PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :		(11) International Publication Number:
C12Q 1/70, 1/68, C12P 19/34 C12N 15/00, C12Q 1/02	A1	WO 90/14443
		(43) International Publication Date: 29 November 1990 (29.11.90)

(21) International Application Number: PCT/US90/02890

(22) International Filing Date: 15 May 1990 (15.05.90)

(30) Priority data:

353,235	16 May 1989 (16.05.89)	US
410,735	20 September 1989 (20.09.89)	US
446,333	4 December 1989 (04.12.89)	US

(71)(72) Applicant and Inventor: HUSE, William, D. [US/US]; 471 Avenida Primavera, Del Mar, CA 92014 (US).

(74) Agents: PERRYMAN, David et al.; Pretty, Schroeder, Brueggemann & Clark, 444 South Flower Street, Suite 2000, Los Angeles, CA 90071 (US).

(81) Designated States: AT, AT (European patent), AU, BB, BE (European patent), BF (OAPI patent), BG, BJ (OAPI patent), BR, CA, CF (OAPI patent), CG (OAPI patent), CH, CH (European patent), CM (OAPI patent), DE*, DE (European patent)*, DK, DK (European patent), ES (European patent), FI, FR (European patent), GA (OAPI patent), GB, GB (European patent), HU, IT (European patent), JP, KP, KR, LK, LU, LU (European patent), MC, MG, ML (OAPI patent), MR (OAPI patent), MW, NL, NL (European patent), NO, RO, SD, SE, SE (European patent), SN (OAPI patent), SU, TD (OAPI patent), TG (OAPI patent), US.

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: CO-EXPRESSION OF HETEROMERIC RECEPTORS

(57) Abstract

The invention provides a composition of matter comprising a plurality of procaryotic cells containing diverse combinations of first and second DNA sequences encoding first and second polypeptides that can be expressed and which form heteromeric receptors and at least one of the plurality of procaryotic cells expressing heteromer exhibiting binding activity towards a preselected molecule.

* See back of page

DESIGNATIONS OF "DE"

Until further notice, any designation of "DE" in any international application whose international filing date is prior to October 3, 1990, shall have effect in the territory of the Federal Republic of Germany with the exception of the territory of the former German Democratic Republic.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MC	Monaco
AU	Australia	FI	Finland	MG	Madagascar
BB	Barbados	FR	France	ML	Mali
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Fasso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GR	Greece	NL	Netherlands
BJ	Benin	HU	Hungary	NO	Norway
BR	Brazil	IT	Italy	RO	Romania
CA	Canada	JP	Japan	SD	Sudan
CF	Central African Republic	KP	Democratic People's Republic of Korea	SE	Sweden
CG	Congo			SN	Senegal
CH	Switzerland	KR	Republic of Korea	SU	Soviet Union
CM	Cameroon	LI	Liechtenstein	TD	Chad
DE	Germany, Federal Republic of	LK	Sri Lanka	TG	Togo
DK	Denmark	LU	Luxembourg	US	United States of America

CO-EXPRESSION OF HETEROMERIC RECEPTORS

BACKGROUND OF THE INVENTION

Many biologically important molecules are proteins, which are composed of linear arrays of amino acid subunits. Proteins can function as enzymes, antibodies or structural proteins, among other things. Proteins whose function is binding other protein, or non-protein molecules, and thereby effect a chemical reaction are termed receptors.

When expressed in a living cell the functional characteristics of proteins are determined by the sequence of their amino acids that are, in turn, encoded by DNA sequences termed genes. While many proteins are single molecules encoded by a single gene, other proteins are composed of two or more separate polypeptides which associate spatially to form an active protein, each polypeptide being encoded by a separate gene. Such proteins are termed heteromers. Where such proteins function as receptors, they are thus heteromeric receptors.

A particular category of protein, as defined by either its characteristic structure or function, exhibits variations in its function, which reflect differences in the particular amino acid sequence. For example, in color vision the receptors for the three different primary colors are the three different rhodopsin molecules which are structurally related but functionally different. Structural differences can also be important in diseases. For example, the hemoglobin of most healthy people and the hemoglobin of individuals with sickle cell anemia differ by a single amino acid. Some categories of proteins in fact exhibit immense variability. Such variability is important because of the particular function of the protein.

Many proteins have multiple structural and functional

domains. In some cases the two different types of domains can coincide. Antibodies are an example of a category of proteins with two well defined structural and functional domains which coincide. One of these domains which bind antigen is functionally diverse and the other, the effector domain, is function restricted. Antibodies are protein comprising four associated polypeptides, two so-called heavy chains, and two light Chains. The four polypeptides associate to form a structure which can be thought of as resembling a "Y", with the tip of the two arms being binding sites which are able to selectively recognize and bind to molecules called antigens, which the body recognizes as foreign. The binding site of the heavy chain is termed VH, while the binding site of the light chain is termed VL. Each arm of the "Y" is called a Fab fragment because it contains the antigen binding functional domain. Such binding is important in order to effect the removal of deleterious foreign materials, for example viruses or bacteria. Because of the vast array of different antigens which an organism may encounter, a vast array of different antibodies are necessary. Such an array, or repertoire, is achieved by an individual having many genes encoding portions of the Vh and VL binding regions. In cells of the immune system, random combinations of these various VH and VL encoding genes can randomly associate in order to allow the expression of upwards of 10 7 different antibody molecules. This possibility arises because the VL and the VH structural domains are smaller than the binding functional domain which is shared between these two structural domains. When a great diversity of functionality can result from the combination of structural domains, the specific function of the combination of any two specific structural domains is not predictable. Therefore, there has been a longstanding problem in protein engineering that combinations of structural domains of proteins which results in predictable function can only generate limited functional diversity, whereas combination

of structural domains which generate the diverse functions are usually unpredictable. Therefore, unpredictability has hampered the construction of protein molecules with highly diverse potential functions. One approach to this problem has been to attempt to increase the predictability of protein design by the rational design of proteins using 3D protein structures and computer algorithms. This approach has not been generally successful. A radically different approach to dealing with the unpredictability would be to construct a very large number of proteins each of which potentially have a desired function. When the unpredictability is matched by the number of potentially correct polypeptides which can be constructed and assayed for a desired function then the problem of unpredictability can be overcome. This has fundamental implications for gene cloning and the design of proteins with predetermined properties.

In the last 15 years, method have been developed in order to produce by expression polypeptide-encoding genes in bacteria, or other cells. This process, which is termed gene cloning, provides the tremendous advantage of allowing the production of large amounts of a particular protein. Genes must be cloned on the basis of their sequence structure or on the basis of the function of the expressed protein. However, until recently it has only been possible to identify single gene clones from a gene library in E. coli, when the cloned genes are to be identified by the function of the expressed protein. Thus it has not been possible to, for example, reproduce the variety of different forms of functions in E. coli. Moreover, even hybridoma technology, which results in large amounts of a single antibody species, suffers from the inability to recreate the vast repertoire of antibody species which can be made even by a single organism, much less those which could be generated within or between species. The ability to generate and screen large repertoires of heteromers in

vitro would potentially allow the selection of particular heteromers having a particular desired function.

There thus exists a long-felt need for a method which can produce vast repertoires of heteromers composed of a plurality of polypeptides each encoded by separate DNA sequences. The present invention satisfies this need and provides related advantages as well.

SUMMARY OF THE INVENTION

The invention provides a composition of matter comprising a plurality of procaryotic cells containing diverse combinations of first and second DNA sequences encoding first and second polypeptides that can be expressed and which form heteromeric receptors and at least one of the plurality of procaryotic cells expressing a heteromer exhibiting binding activity towards a preselected molecule.

The invention further provides a kit for the preparation of vectors useful for the coexpression of two or more DNA sequences, comprising two vectors, a first vector having a first combining site on a defined side of a cloning site which defines orientation and a second vector with a second combining site and a cloning site of orientation asymmetric to that of the first vector, wherein one or both of the vectors contains a promoter for expressing polypeptides which form heteromeric receptors encoded by DNA sequences inserted in the cloning sites.

The invention still further provides a method of constructing a diverse population of vectors having first and second DNA sequences encoding first and second polypeptides which associate to form heteromeric receptors, comprising the steps of

(a) operationally linking a diverse population

of first dna sequences encoding the first polypeptides to a first vector having a c mbining site and a cloning site in a defined orientation;

(b) operationally linking a diverse population of second dna sequences encoding the second polypeptides to a second vector having a combining site compatible with the combining site on the first vector and a cloning site in an asymmetric orientation to that of the first vector;

(c) combining the vector products of step (a) with the vector products of step (b) under conditions to permit their combination into a combined vector having the first and second dna sequences operationally linked thereon. The combining can be accomplished for example, by restriction endonuclease cleavage of the vectors of step (a) and (b) and combining the cleaved vectors of step (a) and (b) with DNA ligase or combining by Flp recombinase.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a schematic diagram of the light chain vector (lambda LC1), the heavy chain vector (lambda Hc2) and the combinatorial vector.

Figure 2 shows nucleotide sequences of the synthetic oligonucleotides inserted into lambda Zap II to create the (A) light chain vector (lambda Lc1) and (B) heavy chain vector (lambda Hc2) of Figure 1.

Figure 3 shows autoradiographs of library screens for the combinatorial (A and B), the heavy chain (E and F) and for the light chain (G and H) libraries. Filter C and D represent the cored positive from a primary filter A.

Figure 4 shows the specificity of antigen binding by competitive inhibition.

Figure 5 is a schematic diagram representing the

plasmids which can be excised from the combinatorial vector.

Figure 6 shows the characterization of an antigen binding protein derived from the combinatorial library.

Figure 7 shows the construction of a vector system for a combinatorial vector using the Flp recognition sequence as the combining site.

DETAILED DESCRIPTION OF THE INVENTION

As used herein "diverse combinations" means that a substantial number of the possible nucleic acids encoding the first polypeptide are combined with a substantial number of the possible nucleic acids encoding the second polypeptide. Thus, a substantial number of the possible combinations are represented.

As used herein "heteromeric receptors" means a polypeptide comprised of at least two polypeptides, at least one of which is encoded on a different DNA. Thus, heteromer is composed of two or more polypeptides which associate and exhibit a common function. Receptor refers toi a polypeptide which is capable of binding any ligand. Therefore, receptor also includes a protein which when bound to its ligand can affect a second process. Examples of heteromeric receptors which can be formed include antibodies, T-cell receptors, integrins, hormone receptors and transmitter receptors.

As used herein "binding activity" means the heteromer exhibits an affinity for a molecule. This affinity can be specific for the molecule and can be used, for example, to detect or affect a function on the molecule.

As used herein "preselected molecule" means a

particular molecule to which binding activity is desired. Since practically any molecule can be bound, this molecule is selected from the gr up of all possible molecules. Specific heteromers can be created which specifically bind this molecule and allow for detection or to affect the molecule's function.

As used herein "first and second polypeptides which can associate" means the polypeptides encoded by the first and second nucleotide sequences are chemically or physically attracted to each other and form a heteromer.

As used herein "combining site" means a nucleotide sequence which can be cleaved and joined with another nucleotide sequence. Such cleavage and joining results in a nucleic acid having both sequences in proper orientation to allow translation of the desired polypeptide.

As used herein "asymmetric" means a non-identical or a non-correspondence in form, size or arrangement of parts on opposite sides of a boundary such as a dividing line or around an axis. For example, the arrangement of 2 different restriction sites with respect to the 5' and 3' ends of a DNA sequence are asymmetric if they are arranged in one vector in the opposite orientation as that for a second vector.

As used herein "transfect" or "transform" refers to introducing nucleic acids into a living cell such that the nucleic acid is fully separated from extracellular fluids by a lipid membrane.

As used herein as it relates to combinatorial gene expression, the term "in vitro" refers to performing the process in a system in which a particular expression does not naturally occur, thus, in vitro can refer both to expression in procaryotic cells as well as eucaryotic

cells, provided the latter does not naturally express the gene combination.

The invention provides a composition of matter comprising a plurality of procaryotic cells containing diverse combinations of first and second DNA sequences encoding first and second polypeptides that can be expressed and which form heteromeric receptors and at least one of the plurality of procaryotic cells expressing a heteromer exhibiting binding activity towards a preselected molecule.

The procaryotic cells are preferably E. coli, however any suitable procaryotic cell can be utilized. Suitable alternative cells would be selected by reviewing the literature to determine which vector and cells could be adapted by the methods taught herein. Therefore, alternative cells require compatible vectors capable of expressing first and second DNA sequences in the selected host cell. Alternatively, eucaryotic cells could be used. Such use would simply require substituting eucaryotic control and expression elements which function in a compatible eucaryotic host. Therefore, for procaryotic and eucaryotic systems, compatibility means that the vector/host combination contains all necessary signals and factors to perform the desired function.

For this invention, including cells, vectors, and methods utilizing the vectors, the first and second DNA sequences which encode functional portions of heteromeric receptors can for example be antibodies, T cell receptors, integrins, hormone receptors and transmitter receptors. Thus, the first and second DNA sequences can encode functional portions of the variable heavy and variable light chains of an antibody including Fab, F'ab and the like. In fact, any heteromer which is formed from a diverse combination or repertoire of alternative coding

sequences can be made by the methods of this invention. For example, specific hormone and transmitter receptors can be made by combination of alpha and beta subunits. Thus, the invention is easily applicable to any later discovered alternative-type, diverse combination heteromers.

The invention also provides a composition of matter comprising a plurality of procaryotic cells containing various combinations of diverse first and second DNA sequences encoding first and second polypeptides which can associate to form heteromeric receptors exhibiting binding activity towards preselected molecules, the diversity of first DNA sequence being greater than about 100 different sequences and the diversity of the second DNA sequence being greater than about 1000 different sequences. The invention is effective with such diversity since the upper limit is greater than a billion combinations.

The invention further provides a kit for the preparation of vectors useful for the coexpression of two or more DNA sequences, comprising two vectors, a first vector having a first combining site on a defined side of a cloning site which defines orientation and a second vector with a second combining site and a cloning site of orientation asymmetric to that of the first vector, wherein one or both of the vectors contains a promoter for expressing polypeptides which form heteromeric receptors encoded by DNA sequences inserted in the cloning sites. The vectors can be in a virus. Suitable virus can include mammalian as well as bacteriophages. One would apply the teachings set forth herein to utilize such vectors. Alternatively, the vectors can be a plasmid.

The first and second combining sites of the vectors of the invention are of many possible types. The specific sites utilized herein are EcoRI-EcoRI, and NotI-NotI and the specific cloning site was selected from the group

consisting of XhoI-SpeI, SacI-XbaI, and SacI-SpeI. Additionally, the first and second combining sites can be site specific recombination sites, especially Flp recombination sites. Alternative sites can be practiced based on the disclosure of this invention.

The invention also provides a vector, capable of expressing a heteromer exhibiting binding activity towards a preselected molecule when combined with a second vector, having a first combining site on a defined side of a cloning site which defines orientation and which can be combined with a second vector with a second combining site and a cloning site of orientation asymmetric to that of the first vector, wherein one or both of the vectors contains a promoter for expressing polypeptides which form heteromers encoded by DNA sequences inserted in the cloning sites.

The invention still further provides a cloning system for the coexpression of two DNA sequences encoding polypeptides which associate to form a heteromer, comprising a set of uniform first vectors having a diverse population of first DNA sequences and a set of uniform second vectors having a diverse population of second DNA sequences, the first and second vectors having compatible combining sites so as to allow the operational combination of the first and second DNA sequences.

The invention also provides a plurality of expression vectors containing a plurality of possible first and second DNA sequences, wherein each of the expression vectors has operationally linked thereon a first DNA sequence and a second DNA sequence, and wherein substantially each of the vectors contains a different combination of first and second DNA sequence.

The invention still further provides a method of

constructing a diverse population of vectors having first and second DNA sequences encoding first and second polypeptides which associate to form heteromeric receptors, comprising the steps of

(a) operationally linking a diverse population of first DNA sequences encoding the first polypeptides to a first vector having a combining site and a cloning site in a defined orientation;

(b) operationally linking a diverse population of second DNA sequences encoding the second polypeptides to a second vector having a combining site compatible with the combining site on the first vector and a cloning site in an asymmetric orientation to that of the first vector;

(c) combining the vector products of step (a) with the vector products of step (b) under conditions to permit their combination into a combined vector having the first and second DNA sequences operationally linked thereon. The combining can be accomplished for example, by restriction endonuclease cleavage of the vectors of step (a) and (b) and combining the cleaved vectors of step (a) and (b) with DNA ligase or combining by Flp recombinase.

A method of selecting a procaryotic cell which expresses a heteromer specific for a preselected molecule is also provided. The method comprises randomly combining first vectors having a diverse population of DNA sequences encoding polypeptides with second vectors having different diverse populations of DNA sequences which encode polypeptides and which form heteromeric receptors with the polypeptides encoded by the first vector, transfecting a sufficient number of the randomly combined sequences into the procaryotic cells, screening the cells to determine the cell expressing a heteromer specific for the preselected molecule. In this method the combining can be accomplished with restriction endonuclease cleavage of the first and second vectors and ligating the cleaved first and second vectors or utilizing Flp recombinase. Additionally, the

number of randomly combined sequences can be sufficiently equivalent to the possible combinations of the populations f the first and second DNAs in order to reasonably assure obtaining the desired heteromer.

Finally, a method is provided for identifying functional heteromeric receptors composed of a plurality of polypeptides, comprising coexpressing random combinations of first and second DNA homologs which encode polypeptides which associate to form heteromeric receptors so as to form a diverse population of the first and second DNA homologs, the diversity being at least enough that at least one heteromer formed by the polypeptides resulting from the coexpression has a desired functional property and restricted so that the heteromeric receptors can be screened for a predetermined function.

In the methods utilized herein, random combination in vitro or in vivo can be accomplished using two expression vectors distinguished from one another by the location of an endonuclease recognition site common to both. Preferably the vectors are linear double stranded DNA, such as a lambda Zap™ derived vector as described herein which are symmetric with respect to the protein expression elements. Preferably, in one of the vectors the recognition site is located 5' terminal to the coding sequence of at least one of the complementary determining regions (CDR's). In the second vector the recognition site is located 3' to at least one of the CDR's. For example, the recognition site in one vector can be located between a ribosome binding site and a RNA polymerase promoter site and in the second vector the restriction is located 3' to a cloning site.

The recognition site can be a restriction endonuclease recognition site, a recombinase recognition site such as a Flp site, or other equivalent site. In one preferred

embodiment of the invention, each of the vectors defines a nucleotide sequence coding for a ribosome binding site and a leader, the sequence being located between the promoter and the polylinker, but downstream (3' terminal from a shared restriction site if that site is between the promoter and the polylinker). Also preferred are vectors containing a stop codon downstream from the polylinker. The first and/or second vector can also define a nucleotide sequence coding for a polypeptide which can function as a tag. Examples of such a tag include (1) a short peptide sequence, (2) a sequence that encodes a protein, which binds to a receptor such as another predetermined antibody or protein G, such as a CH1 domain of an antibody, (3) a protein that can function as an enzyme (such as beta-galactosidase or alkaline phosphatase) or (4) a phage coat protein that causes the phage to become attached to the coat of the phage. The tag sequence is typically downstream from the polylinker but upstream of any stop codon that may be present. In the preferred embodiments, the vectors contain selectable markers such that the presence of a portion of that vector, i.e. a particular lambda arm, can be selected for or selected against.

Typical selectable markers are well known to those skilled in the art. Examples of such markers are antibiotic resistance genes, genetically selectable markers, suppressible mutations, such as amber mutations, and the like. The selectable markers are typically located upstream and/or downstream of the promoter or polylinker. In preferred embodiments, one selectable marker is located upstream of the promoter on the first vector containing the VH-coding (variable heavy chain-coding) DNA sequences. A second selectable marker is located on the other side of the combination site on the vector containing the VL-coding (variable light chain-coding) DNA sequences. This second selectable marker may be the same or different from the first as long as when the VH-coding vectors and the VL-

coding vectors are randomly combined at the combining site the resulting vectors containing both VH and VL can be selected preferentially.

Typically the polylinker is a nucleotide sequence that defines one or more, preferably at least two, restriction sites. The polylinker restriction sites are oriented to permit ligation of VH- or VL-coding DNA homologs into the vectors in the same reading frame at the leader, tag, linker, tag, or stop codon sequence present.

Random combination is accomplished by ligating VH-coding DNA homologs into the first vector, typically at a restriction site or sites within the polylinker. Similarly, VL-coding DNA homologs are ligated into the second vector, thereby creating two diverse populations of vectors. It does not matter which type of DNA homolog, i.e., VH or VL, is ligated to which vector, but it is preferred, for example, that all VH coding DNA homologs are ligated to either the first or second vector, and all of the VL-coding DNA homologs are ligated to the other of the first or second vector. The members of both populations are combined at the combination site. In a preferred embodiment where the combination site is a restriction site and the members of both populations are then cleaved with an appropriate restriction endonuclease. The resulting products are two diverse populations of restriction fragments where the members of one have cohesive termini complementary to the cohesive termini of the members of the other.

The following examples are intended to illustrate but not limit the invention. While they are typical of those that might be used, other procedures known to those skilled in the art may be alternatively employed.

EXAMPLE I

VECTOR CONSTRUCTION

The vectors for expression of V_H, V_L, Fv (fragment of the variable region), and Fab sequences are diagrammed in Figures 1 and 2. They were constructed by a modification of lambda Zap II, (Stratagene, La Jolla, CA); Short et al., Nucleic Acids Res., 16:7583 (1988) which is incorporated herein by reference, in which we inserted synthetic oligonucleotides into the multiple cloning site. The methods described here and below are known to one skilled in the art and are described in detail in Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, 1982 and Ausubel et al., and Current Protocols on Molecular Biology, John Wiley and Sons, 1987, both of which are incorporated herein by reference. The vectors were designed to be asymmetric with respect to the Not I and Eco RI restriction sites that flank the cloning and expression sequences. This asymmetry in the placement of restriction sites in a linear vector such as bacteriophage allows a library expressing light chains to be combined with one expressing heavy chains in order to construct combinatorial Fab expression libraries.

The lambda Lc 1 vector was constructed for the cloning of PCR amplified products of mRNA that code for light chain protein, as described in Example II, by inserting the nucleotide sequence shown in figure 2A into the Sac I and Xho I sites of lambda Zap II. The vector was prepared by digesting 10 μg of lambda arms from the Uni-Zap™ XR Vector Kit (Stratagene, La Jolla, CA) with Sac I. The sequence shown in Figure 2A was constructed from overlapping synthetic oligonucleotides and cloned into the above Sac I digested arms as follows. Oligonucleotides L1 through L5 and L7 - L9 (L1, L2, L3, L4, L5, L7, L8 and L9) (shown in Table 1) were kinased by adding 1 μl of each oligonucleotide (0.1 μg/μl) and 20 units of T_4

polynucleotide kinase (BRL, Gaithersburg, MD) to a solution containing 70 mM Tris HCL at pH 7.6, 0.1 M KCl, 10 mM $MgCl_2$, 5 mM DTT, 1 mM adenosine triphosphate (ATP), 10 mM 2 ME, 500 micrograms per ml of BSA. The solution was maintained at 37°C for 30 minutes and the reaction stopped by maintaining the solution at 65°C for 10 minutes. The two end oligonucleotides L6 and L10 were added to the above kinasing reaction solution together with 1/10 volume of a solution containing 20 mM Tris-HCL at pH 7.4, 2.0 mM $MgCl_2$ and 50.0 mM NaCl. This solution was heated to 70°C for 5 minutes and allowed to cool slowly to room temperature. During this time period all oligonucleotides annealed to form the double stranded synthetic DNA insert shown in Figure 2A. The annealed oligonucleotides were covalently linked to each other by adding 40 μl of the above reaction to a solution containing 66 mM Tris-HCL at pH 7.6, 6.6 mM $MgCl_2$, 1 mM DTT, 1 mM ATP and 10 units of T4 DNA ligase (BRL, Gaithersburg, MD). This solution was maintained at 25°C for 30 minutes and then the T4 DNA ligase was inactivated by heating the solution at 65°C for 10 minutes. The unphosphorylated ends of the resultant oligonuleotides were kinased by mixing 52 μl of the above reaction, 4 μl of a solution containing 10 mM ATP and 5 units of T4 polynucleotide kinase. This solution was maintained at 37°C for 30 minutes and then the T4 polynucleotide kinase was inactivated by heating the solution at 65°C for 10 minutes. The phosphorylated synthetic DNA insert was ligated directly into the above prepared lambda Zap II vector arms.

TABLE 1

L1	TGAATTCTAAACTAGTCGCCAAGGAGACAG
L2	TCATAATGAAATACCTATTGCCTACGGCAG
L3	CCGCTGGATTGTTATTACTCGCTGCCCAAC
L4	CAGCCATGGCCGAGCTCGTCAGTTCTAGAG
L5	TTAAGCGGCCGCAA
L6	TCGATTGCGGCCGCTTAACTCTAGAACTGACGA
L7	GCTCGGCCATGGCTGGTTGGGCAGCGAGTA
L8	ATAACAATCCAGCGGCTGCCGTAGGCAATA
L9	GGTATTTCATTATGACTGTCTCCTTGGCGA
L10	CTAGTTTAGAATTCAAGCT

TABLE 2

H1	GGCCGCAAATTCTATTTCAAGGAGACAGTC
H2	ATAATGAAATACCTATTGCCTACGGCAGCC
H3	GCTGGATTGTTATTACTCGCTGCCCAACC
H4	AGCCATGGCCCAGGTGAAACTGCTCGAGA
H5	TTTCTAGACTAGTTACCCGTACGACGTTCC
H6	GGACTACGGTTCTTAATAGAATTCG
H7	TCGACGAATTCTATTA
H8	AGAACCGTAGTCCGGAACGTCGTACGGG
H9	TAACTAGTCTAGAAATCTCGAGCAGTTTC
H10	ACCTGGGCCATGGCTCCTTGGGCAGCGAGT
H11	AATAACAATCCAGCGGCTGCCGTAGGCAA
H12	TAGGTATTTCATTATGACTGTCTCCTT
H13	GAAATAGAATTTGC

The lambda Hc 2 vector was constructed for cloning PCR amplified products coding for heavy chain Fd sequences, as described in Example II, by inserting the nucleotide sequence shown in Figure 2B into the Not I and Xho I sites of lambda Zap II. As with the light chain vector, the heavy chain vector was prepared by digesting lambda arms from the Uni-Zap™ XR Vector Kit (Stratagene, La Jolla, CA) with Not I. This was accomplished by digestion of 10 μg of

vector in 100 μl reaction buffer for 1 hour at 37°C, after digestion the DNA was extracted, precipitated and dried as above. The inserted sequence shown in Figure 2B was constructed from the overlapping synthetic oligonucleotides H1-H13 depicted in Table 2 as outlined above. Correctly constructed vectors were confirmed by DNA sequence analysis as described below.

The sequence of the oligonucleotides described above include elements for construction, expression, and secretion of Fab fragments. These oligonucleotides introduce the asymmetric Not I and Eco RI restriction sites; a leader peptide for the bacterial pel B gene, which has previously been successfully used in E. coli to secrete Fab fragments, Better et al., Science, 240:1041 (1988); Skerra and Pluckthun, Science, 240:1038 (1988), both of which are incorporated herein by reference, a ribosome binding site at the optimal distance for expression of the cloned sequence; cloning sites for either the light or heavy chain PCR product; and, in lambda Hc 2, a decapeptide tag at the carboxyl terminus of the expressed heavy chain protein fragment. The sequence of the decapeptide tag was useful because of the availability of monoclonal antibodies to this peptide that were used for immunoaffinity purification of fusion proteins, Field et al. Mol. Cell Biol., 8:2159 (1988), which is incorporated herein by reference. The vectors were characterized by restriction digest analysis and DNA sequencing, Sanger et al., Proc. Natl. Acad. Sci., USA, 74:5463-5467 (1977), which is incorporated herein by reference and using AMV Reverse Transcriptase ^{35}S-ATP Sequencing Kit (Stratagene, La Jolla, CA).

EXAMPLE II

Isolation of mRNA and PCR
Amplification of Antibody Fragments

The initial Fab expression library was constructed from mRNA isolated from a mouse that had been immunized with the KLH-coupled p-nitrophenyl phosphonamidate antigen 1 (NPN). NPN was coupled to keyhole limpet hemocyanin (KLH) using the techniques described in Antibodies: A Laboratory Manual, Harlow and Lowe, eds., Cold Spring Harbor, New York (1988), which is incorporated herein by reference. Briefly, 10.0 milligrams (mg) of keyhole limpet hemocyanin and 0.5 mg of NPN with a glutaryl spacer arm N-hydroxysuccinimide linker appendages. Coupling was performed as in Jonda et al., Science, 241:1188 (1988), which is incorporated herein by reference. The unbound NPN was removed by gel filtration chromatography through Sephadex G-25.

The KLH-NPN conjugate was prepared for injection into mice by adding 100 μg of the conjugate to 250 μl of phosphate buffered saline (PBS). An equal volume of complete Freund's adjuvant was added and emulsified the entire solution for 5 minutes. A 129 G_{IX+} mouse was injected with 300 μl of the emulsion. Injections were given subcutaneously at several sites using a 21 gauge needle. A second immunization with KLH-NPN was given two weeks later. This injection was prepared as follows: 50 μg of KLH-NPN were diluted in 250 μL of PBS and an equal volume of alum was mixed with the KLH-NPN solution. The mouse was injected intraperitoneally with 500 μl of the solution using a 23 gauge needle. One month later the mice were given a final injection of 50 μg of the KLH-NPN conjugate diluted to 200 μL in PBS. This injection was given intravenously in the lateral tail vein using a 30 gauge needle. Five days after this final injection the mice were

sacrificed and total cellular RNA was isolated from their spleens.

Total RNA was isolated from the spleen of a single mouse immunized as described above by the method of Chomczynski and Sacchi, Anal. Biochem., 162:156-159 (1987), which is incorporated herein by reference. Briefly, immediately after removing the spleen from the immunized mouse, the tissue was homogenized in 10 ml of a denaturing solution containing 4.0 M guanine isothiocyanate, 0.25 M sodium citrate at pH 7.0, and 0.1 M 2-mercaptoethanol using a glass homogenizer. One ml of sodium acetate at a concentration of 2 M at pH 4.0 was mixed with the homogenized spleen. One ml of saturated phenol was also mixed with the denaturing solution containing the homogenized spleen. Two ml of a chloroform:isoamyl alcohol (24:1 v/v) mixture was added to this homogenate. The homogenate was mixed vigorously for ten seconds and maintained on ice for 15 minutes. The homogenate was then transferred to a thick-walled 50 ml polypropylene centrifuge tube (Fisher Scientific Company, Pittsburgh, PA). The solution was centrifuged at 10,000 x g for 20 minutes at 4°C. The upper RNA-containing aqueous layer was transferred to a fresh 50 ml polypropylene centrifuge tube and mixed with an equal volume of isopropyl alcohol. This solution was maintained at -20°C for at least one hour to precipitate the RNA. The solution containing the precipitated RNA was centrifuged at 10,000 x g for twenty minutes at 4°C. The pelleted total cellular RNA was collected and dissolved in 3 ml of the denaturing solution described above. Three ml of isopropyl alcohol was added to the resuspended total cellular RNA and vigorously mixed. This solution was maintained at -20°C for at least 1 hour to precipitate the RNA. The solution containing the precipitated RNA was centrifuged at 10,000 x g for ten minutes at 4°C. The pelleted RNA was washed once with a solution containing 75% ethanol. The pelleted RNA was

dried under vacuum for 15 minutes and then resuspended in dimethyl pyrocarbonate (DEPC) treated (DEPC-H_2O) H_2O.

Poly A^+ RNA for use in first strand cDNA synthesis was prepared from the above isolated total RNA using methods described by Aviv and Leder, Proc. Natl. Acad. Sci., USA, 69:1408-1412 (1972), which is incorporated herein by reference. Briefly, one half of the total RNA isolated from a single immunized mouse spleen prepared as described above was resuspended in one ml of DEPC-treated dH_2O and maintained at 65°C for five minutes. One ml of 2x high salt loading buffer (100 mM Tris-HCL at pH 7.5, 1 M sodium chloride, 2.0 mM disodium ethylene diamine tetraacetic acid (EDTA) at pH 8.0, and 0.2% sodium dodecyl sulfate (SDS)) was added to the resuspended RNA and the mixture was allowed to cool to room temperature. The mixture was then applied to an oligo-dT (Collaborative Research Type 2 or Type 3) column that was previously prepared by washing the oligo-dT with a solution containing 0.1 M sodium hydroxide and 5 mM EDTA and then equilibrating the column with DEPC-treated dH_2O. The eluate was collected in a sterile polypropylene tube and reapplied to the same column after heating the eluate for 5 minutes at 65°C. The oligo dT column was then washed with 2 ml of high salt loading buffer consisting of 50 mM Tris-HCL at pH 7.5, 500 mM sodium chloride, 1 mM EDTA at pH 8.0 and 0.1% SDS. The oligo dT column was then washed with 2 ml of 1 X medium salt buffer (50 mM Tris-HCL at pH 7.5, 100 mM sodium chloride, 1 mM EDTA at pH 8.0 and 0.1% SDS). The mRNA was eluted with 1 ml of buffer consisting of 10 mM Tris-HCL at pH 7.5, 1 mM EDTA at pH 8.0 and 0.05% SDS. The messenger RNA was purified by extracting this solution with phenol/chloroform followed by a single extraction with 100% chloroform, ethanol precipitated and resuspended in DEPC treated dH_2O.

In preparation for PCR amplification, mRNA was used as

a template for cDNA synthesis. In a typical 250 μl transcription reaction mixture, 5-10μg of spleen mRNA in water was first annealed with 500 ng (0.5 pmol) of either the 3' V_H primer (primer 12, Table 3) or the 3' V_L primer (primer 9, Table 4) at 65°C for 5 minutes. Subsequently, the mixture was adjusted to contain 0.8 mM dATP, 0.8 mM dCTP, 0.8 mM dGTP, 0.8 mM dTTP, 100 mM Tris-HCL (pH 8.6), 10 mM $MgCl_2$, 40 mM KCl, and 20 mM 2-ME. Moloney-Murine Leukemia Virus (Stratagene, La Jolla, CA) Reverse transcriptase, 26 units, was added and the solution was incubated for 1 hour at 40°C. The resultant first strand cDNA was phenol extracted, ethanol precipitated and then used in the polymerase chain reaction (PCR) procedures described below for amplification of heavy and light chain sequences.

Primers used for amplification of heavy chain Fd fragments for construction of the lambda Hc 2 library is shown in Table 3. Amplification was performed in eight separate reactions, as described by Saiki et al., Science, 239:487-491 (1988), which is incorporated herein by reference, each reaction containing one of the 5' primers (primers 2 to 9) and one of the 3' primers (primer 12) listed in Table 3. The remaining 5' primers were used for amplification in a single reaction are either a degenerate primer (primer 1) or a primer that incorporates inosine at four degenerate positions (primer 10). The remaining 3' primer (primer 11) was used to construct Fv fragments. The underlined portion of the 5' primers incorporates an Xho I site and that of the 3' primer an Spe I restriction site for cloning the amplified fragments into a lambda phage vector in a predetermined reading frame for expression.

TABLE 3

HEAVY CHAIN PRIMERS

1) 5'- AGGT CC G G T / GA A T A A CT CTCGAGTC GG - 3'

2) 5' - AGGTCCAGCTGCTCGAGTCTGG - 3'

3) 5' - AGGTCCAGCTGCTCGAGTCAGG - 3'

4) 5' - AGGTCCAGCTTCTCGAGTCTGG - 3'

5) 5' - AGGTCCAGCTTCTCGAGTCAGG - 3'

6) 5' - AGGTCCAACTGCTCGAGTCTGG - 3'

7) 5' - AGGTCCAACTGCTCGAGTCAGG - 3'

8) 5' - AGGTCCAACTTCTCGAGTCTGG - 3'

9) 5' - AGGTCCAACTTCTCGAGTCAGG - 3'

10) 5' - AGGTIIAICTICTCGAGTC T GG - 3' / A

11) 5' - CTATTAACTAGTAACGGTAACAGT - GGTGCCTTGCCCCA - 3'

12) 5' - AGGCTTACTAGTACAATCCCTGG - GCACAAT - 3'

Primers used for amplification of mouse kappa light chain sequences for construction of the lambda Lc 1 library is Fab's are shown in Table 4. These primers were chosen to contain restriction sites which were compatible with vector and not present in the conserved sequences of the mouse light chain mRNA. Amplification was performed as described by Saiki et al., Supra, in five separate reactions, each containing one of the 5' primers (primers 3 to 7) and one of the 3' primers (primer 9) listed in Table 4. The remaining 3' primer (primer 8) was used to construct Fv fragments. The underlined portion of the 5' primers depicts a Sac I restriction site and that of the 3' primers an Xba I restriction site for cloning of the amplified fragments into a lambda phage vector in a

predetermined reading frame for expression.

TABLE 4

LIGHT CHAIN PRIMERS

1) 5' - CCAGTTCCGAGCTCGTTGTGACTCAGGAATCT - 3'
2) 5' - CCAGTTCCGAGCTCGTGTTGACGCAGCCGCCC - 3'
3) 5' - CCAGTTCCGAGCTCGTGCTCACCCAGTCTCCA - 3'
4) 5' - CCAGTTCCGAGCTCCAGATGACCCAGTCTCCA - 3'
5) 5' - CCAGATGTGAGCTCGTGATGACCCAGACTCCA - 3'
6) 5' - CCAGATGTGAGCTCGTCATGACCCAGTCTCCA - 3'
7) 5' - CCAGTTCCGAGCTCGTGATGACACAGTCTCCA - 3'
8) 5' - GCAGCATTCTAGAGTTTCAGCTCCAGCTTGCC - 3'
9) 5' - GCGCCGTCTAGAATTAACACTCATTCCTGTTGAA - 3'

PCR amplification for heavy and light chain fragments was performed in a 100-μl reaction mixture containing the above described products of the reverse transcription reaction ($\approx$5μg of the cDNA-RNA hybrid), 300 nmol of 3' V_H primer (primer 12, Table 1), and one of the 5' V_H primers (primers 2-9, Table 1) for heavy chain amplification, or, 300 nmol of 3' V_L primer (primer 9, Table 2), and one of the 5' V_L primers (primers 3-7, Table 2) for each light chain amplification, a mixture of dNTPs at 200 mM, 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 15 mM $MgCl_2$, 0.1% gelatin, and 2 units of Thermus aquaticus DNA polymerase. The reaction mixture was overlaid with mineral oil and subjected to 40 cycles of amplification. Each amplification cycle involved denaturation at 92°C for 1 minute, annealing at 52°C for 2 minutes, and elongation at 72°C for 1.5 minutes. The amplified samples were extracted twice with phenol/$CHCl_3$ and once with $CHCl_3$, ethanol-precipitated, and stored at -70°C in 10 mM Tris-HCl, pH 7.5/1 mM EDTA.

In preparation for cloning into the lambda Hc 2 or lambda Vc 1 vectors equal volumes (50 μl) of the above, respective, PCR-amplifi d products were mixed, purified by

phenol/$ChCl_3$ extraction, ethanol precipitated and resuspended at 1 μg/μl in 10 mM Tris-HCl, pH 7.5/1 mM EDTA. The mixed products f heavy chain primer PCR amplification were digested at 37°C with Xho I (125 units, Stratagene, La Jolla, CA) and Spe I (10 units, Stratagene, La Jolla, CA) in 2.5 μg/30 μl of buffer containing 150 mM NaCl, 8 mM Tris-HCl (pH 7.5), 6 mM $MgSO_4$, 1 mM dithiothreitol, and bovine serum albumin (200 μg/ml). The mixed products of amplification with light chain primers were digested with 200 units Sac I and 200 units Xba I in 33 mM Tris Acetate pH 7.85, 66 mM K Acetate, 10 mM Mg Acetate, 0.5 mM DTT in 500 μl at 37°C for 1 hour for the light chain amplified products and purified on a 1% agarose gel. After gel electrophoresis of the digested PCR-amplified spleen mRNA the region of the gel containing DNA fragments of 700 base pairs (bp) was excised, electroeluted into a dialysis membrane, ethanol-precipitated, and resuspended in 10 mM Tris-HCl, pH 7.5/1 mM EDTA to a final concentration of 10 ng/μl. These products were used in the library constructions described in Example III.

EXAMPLE III
LIBRARY CONSTRUCTION

A combinatorial library was constructed in two steps. In the first step, separate heavy and light chain libraries were constructed in lambda Hc 2 and lambda Lc 1 vectors, respectively (Figure 1). In the second step, the two resultant libraries were combined at the asymmetric Eco RI sites present in each vector.

For construction of lambda Hc 2 and Lambda Lc 1 libraries, 3 molar equivalence of the gel isolated inserts described in Example II were ligated with 1 molar equivalence of vector arm, as described below overnight at 5°C to lambda Hc 2 or lambda Lc 1, described in Example I. The heavy chain inserts were ligated to lambda Hc 2 arms

previously digested with Xho I and Spe I and dephosphorylated. The light chain inserts were ligated to lambda Lc 1 arms previously digested with Sac I and Xba I and dephosphorylated. Vector arms were prepared using the techniques described in Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, which is incorporated herein by reference. 10 ml of NZCYM broth (10g/l NZ amine, 5g/l yeast extract, 5g/l NaCl, 1g/l casamino acids, 2g/l $MgSO_4$-$7H_2O$, pH 7.5) was inoculated with a single colony of XL1-Blue and incubated overnight with vigorous agitation at 37°C. 1 ml of this culture was used to inoculate four 2 liter flasks containing 500 ml prewarmed (37°C) NZCYM. These four flasks were agitated at 37°C 3 to 4 hours. Each flask then received an inoculation of 10^{10} pfu of the purified recombinant bacteriophage vector prepared in Example I, and was shaken for an additional 3 to 5 hours until lysis of the host was complete. 10 ml of chloroform was added to each flask and incubation continued for another 10 minutes at 37°C. Cultures were treated with 1 μg/ml each DNAsE I and RNAseA for 30 minutes at room temperature. NaCl was added to 1 M final concentration and the cultures were chilled on ice for 1 hour. Debris was removed by centrifugation at 11,000xg for 10 minutes, and polyethylene glycol (PEG 8000) was added to the supernatants to a final concentration of 10% w/v. Bacteriophage precipitated out of the suspensions after 1 hour on ice and was pelleted by centrifugation at 11,000xg for 10 minutes. Phage was resuspended in SM buffer (5.8g/l NaCl, 2g/l $MgSO_4$-7 H_2O, 50ml/l 1 M Tris-Cl pH 7.5, 5ml 2% gelatin) and chloroform extracted to remove cell debris. Solid cesium chloride (CsCl) was added to 0.5g/ml, and the phage suspension was layered onto CsCl step gradients (1.7g/ml, 1.5g/ml, 1.45g/ml, all in SM) and spun at 22,000 rpm for 2 hours at 4°C in a swinging bucket rotor. Banded phage particles were collected and spun in 1.5g/ml CsCl/SM at 38,000 rpm for 24 hours at 4°C. Re-banded phage was again collected, and the suspension was dialysed in 10mM

NaCl, 50mM Tris-Cl pH 8.0, 10mM $MgCl_2$. EDTA pH 8.0 was added to 20mM, pronase was added to 0.5mg/ml, and SDS was added to 0.5%; incubation at 37° for 1 hour was followed by phenol extraction, chloroform extraction, and dialysis overnight in 10mM Tris-Cl pH 8.0, 1mM EDTA pH 8.0. Sodium acetate was added to 0.3 M and the DNA was precipitated with 2 volumes of ethanol. Vector DNA was recovered by centrifugation and resuspended in 10mM Tris-Cl pH 7.6, 1mM EDTA pH 8.0.

To make Hc2 vector arms, 200 μg purified Hc2 DNA was cut with 600 units Xho I in 50mM Tris-Cl pH 8.0, 10mM $MgCl_2$, 50mM NaCl, at 37°C for 1 hour. Cut HC2 DNA was phenol extracted and ethanol precipitated, then re-cut with 600 units of Spe I in 20mM Tris Cl pH 7.4, 5mM $MgCl_2$, 50mM KCl at 37°C for 1 hour. Double-cut Hc2 DNA was phenol extracted and ethanol precipitated. Recovered vector DNA was dephosphorylated with 0.5 units/μg HK phosphatase (Epicenter, Madison, WI) in 30mM Tris Acetate pH 7.85, 30mM KAC, 5mM $CaCl_2$, 0.5mM DTT, and 100 μg/ml BSA, at 30° for 1 hour, followed by 65° for 10 minutes, then phenol extracted, ethanol precipitated and resuspended in 10mM Tris Cl pH 7.5, 1 mM EDTA pH 8.0.

Lc1 vector arms were prepared as above, except that the first digestion was with 600 units of Xba I in 50mM Tris-Cl pH 8.0, 10mM $MgCl_2$, 50mM NaCl, and the second digestion was with 600 units of Sac I in 6mM Tris Cl pH 7.4, 20mM NaCl, 6mM $MgCl_2$, 6mM 2-ME, 0.1 mg/ml BSA. A portion of each ligation mixture (1 μl) was packaged for 2 hours at room temperature using Gigapack Gold packaging extract (Stratagene, La Jolla, CA), and the packaged material was titered and plated on XL1-Blue host cells as described by the manufacturer.

Specifically, serial dilutions of the library were made into a buffer containing 100 mM NaCl, 50 mM Tris-HCL

at pH 7.5 and 10 mM $MgSO_4$. Ten μl of each dilution was added to 200 μl of exponentially growing E. coli cells and maintained at 37°C for 15 minutes to allow the phage to absorb to the bacterial cells. Three ml of top agar consisting of 5 g/L NaCl, 2 g/L of $MgSO_4$, 5 g/L yeast extract, 10 g/L NZ amine (casein hydrolysate) and 0.7% melted, 50C agarose. The phage, the bacteria and the top agar were mixed and then evenly distribute across the surface of a prewarmed bacterial agar plate (g g/L NaCl, 2 g/L $MgSO_4$, 5 g/L yeast extract, 10 g/L NZ amine (casein hydrolysate) and 15 g/L Difco agar. The plates were maintained at 37°C for 12 to 24 hours during which time period the lambda plaques were counted to determine the total number of plaque forming units per ml in the original library.

The lambda Hc 2 primary library contained 1.3×10^6 plaque-forming units (pfu) and has been screened for the expression of the decapeptide tag to determine the percentage of clones expressing Fd sequences. The sequence for this peptide is only in frame for expression after the genes for an Fd (or V_H) fragment have been cloned into the vector. At least 80 percent of the clones in the library express Fd fragments when assayed by immunodetection of the decapeptide tag.

Immunodetection was performed as follows. A volume of the titred library that would yield 20,000 plaques per 150 millimeter plate was added to 600 μl of exponentially growing E. coli cells and maintained at 37°C for 15 minutes to allow the phage to absorb to the bacterial cells. Then 7.5 ml of top agar was admixed to the solution containing the bacterial cells and the absorbed phage and the entire mixture distributed evenly across the surface of a prewarmed bacterial agar plate. This process was repeated for a sufficient number of plates to plate out a total number of plaques at least equal to the library size.

These plates were then maintained at 37°C for 5 hours. The plates were then overlaid with nitrocellulose filters that had been pretreated with a solution containing 10 mM isopropyl-beta-D-thiogalactopyranoside (IPTG) and maintained at 37°C for 4 hours. The orientation of the nitrocellulose filters in relation to the plate were marked by punching a hole with a needle dipped in waterproof ink through the filter and into the bacterial plates at several locations. The nitrocellulose filters were removed with forceps and washed once in a TBST solution containing 20 mM Tris-HCL at pH 7.5, 150 mM NaCl and 0.05% polyoxyethylene soriban monolaurate (Tween-20). A second nitrocellulose filter that had also been soaked in a solution containing 10 mM IPTG was reapplied to the bacterial plates to produce duplicate filters. The filters were further washed in a fresh solution of TBST for 15 minutes. Filters were then placed in a blocking solution consisting of 20 mM Tris-HCL at pH 7.5, 150 mM NaCl and 1% BSA and agitated for 1 hour at room temperature. The nitrocellulose filters were transferred to a fresh blocking solution containing a 1 to 500 dilution of the primary antibody and gently agitated for at least 1 hour at room temperature. After the filters were agitated in the solution containing the primary antibody the filters were washed 3 to 5 times in TBST for 5 minutes each time to remove any of the residual unbound primary antibody. The filters were transferred into a solution containing fresh blocking solution and a 1 to 500 to a 1 to 1,000 dilution of alkaline phosphatase conjugated secondary antibody. The filters were gently agitated in the solution for at least 1 hour at room temperature. The filters were washed 3 to 5 times in a solution of TBST for at least 5 minutes each time to remove any residual unbound secondary antibody. The filters were washed once in a solution containing 20 mM Tris-HCL at pH 7.5 and 150 mM NaCl. The filters were removed from this solution ad the excess moisture blotted from them with filter paper. The color was developed by placing the filter in a solution

containing 100 mM Tris-HCL at pH 8.5, 100 mM NaCl, 5 mM $MgCl_2$, 0.3 mg/ml of nitro Blue Tetrazolium (NBT) and 0.15 mg/ml of 5-bromo-4-chloro-3-indolyl-phosphate (BCIP) for at least 30 minutes at room temperature. The residual color development solution was rinsed from the filter with a solution containing 20 mM Tris-HCL at pH 7.5 and 150 mM NaCl. The filter was then placed in a stop solution consisting of 20 mM Tris-HCL at pH 2.9 and 1 mM EDTA. The development of an intense purple color indicates at positive results. The filters are used to locate the phage plaque that produced the desired protein. That phage plaque is segregated and then grown up for further analysis.

The light chain library was constructed in the same way as the heavy chain and shown to contain 2 x 10^6 members. Plaque screening, with an antibody to mouse kappa chain, indicated that 60 percent of the library contained expressed light chain inserts. This relatively small percentage of inserts probably resulted from incomplete dephosphorylation of the vector after cleavage with Sac I and Xba I.

For construction of the combinatorial library, the above two libraries were used by crossing them at the Eco RI site as follows. DNA was first purified from each library as described above. The light chain library was cleaved with Mlu I restriction endonuclease, the resulting 5' ends were dephosphorylated, and the product was digested with Eco RI. This process cleaved the left arm of the vector into several pieces, but the right arm containing the light chain sequences remained intact. The DNA of heavy chain library was cleaved with Hind III, dephosphorylated, and then cleaved with Eco RI; this process destroyed the right arm, but the left arm containing the heavy chain sequences remained intact. The DNA's so prepared were then mixed and ligated. After

ligation, only clones that resulted from combination of a right arm of light chain-containing clones and left arm of heavy chain-containing clones reconstituted a viable phage. After ligation and packaging, 2.5×10^7 clones were obtained. This is the combinatorial Fab expression library that was screened to identify clones having affinity for NPN as described below in Example IV. For determining the frequency of the phage clones that coexpress the light and heavy chain fragments, duplicate lifts of the combinatorial library for light and heavy expression were screened. In the examination of approximately 500 recombinant phage, approximately 60 percent coexpressed light and heavy chain proteins.

EXAMPLE IV
ANTIGEN BINDING

All three libraries, the light chain, the heavy chain, and Fab were screened to determine whether they contained recombinant phage that expressed antibody fragments binding NPN. In a typical procedure, 30,000 phage were plated and duplicate lifts with nitrocellulose were screened as described in Example III for binding to NPN coupled to ^{125}I-labeled bovine serum albumin (BSA) (Figure 3). Duplicate screens of 90,000 recombinant phage from the light chain library and a similar number from the heavy chain library did not identify any clones that bound the antigen. In contrast, the screen of a similar number of clones from the Fab expression library identified many phage plaques that bound NPN (Figure 5). Briefly, duplicate plaque lifts of Fab (filters A and B), heavy chain (filters E and F), and light chain (filters G and H) expression libraries were screened against ^{125}I-labeled BSA conjugated with NPN at a density of approximately 30,000 plaques per plate. Filters C and D illustrate the duplicate secondary screening of a cored positive from a primary filter A (arrows). BSA was labeled as described in Harlow et al., Supra, which is

inc rporated herein by reference, and coupling reactions were as described in Example II. Standard plaque lift methods were used in screening as described in Example II and in Ausubel et al., Current Protocols in Molecular Biology, John Wiley and Sons, (1987), Supra. Briefly, cells (XLI blue) infected with phage were incubated on 150-mm plates for 4 hours at 37°C, protein expression was induced by overlay with nitrocellulose filters soaked in 1 mM Isopropyl-1-thio-β-D-galactoside (IPTG) and the plates were incubated at 25°C for 8 hours. Duplicate filters were obtained during a second incubation under the same conditions. Filters were then blocked in a solution of 1 percent BSA in phosphate-buffered saline (PBS) for 1 hour before incubation (with rocking) at 25°C for 1 hour with a solution of ^{125}I-labeled BSA (at 0.1 μM) conjugated to NPN (2×10^6 cpm/ml; approximately 15 NPN per BSA molecule), in 1 percent BSA in PBS. Background was reduced by preliminary centrifugation of stock ^{125}I-labeled BSA solution at 100,000g for 15 minutes and preliminary incubation of solutions with plaque lifts from plates containing bacterial infected with a phage having no insert. After labeling, filters were washed repeatedly with PBS containing 0.05 percent Tween 20 before the overnight development of autoradiographs.

This observation indicates that, under conditions where many heavy chains in combination with light chains bind to antigen, heavy or light chains alone do not. Therefore, in the case of NPN, there are many heavy and light chains that only bind antigen when they are combined with specific light and heavy chains, respectively. This result supports our decision to screen large combinatorial Fab expression libraries. To assess our ability to screen large numbers of clones and obtain a more quantitative estimate of the frequency of antigen binding clones in the combinatorial library, we screened one million phage plaques and identified approximately 100 clones that bound

to antigen. For six clones, a region of the plate containing the positive phage plaques and approximately 20 surrounding them was "cored," replated, and screened with duplicate lifts (Figure 3). The expression products of approximately 1 in 20 of the phage specifically bind to antigen. Phage which were believed to be negative on the initial screen did not give positives on replating.

To determine the specificity of the antigen-antibody interaction, antigen-binding was subjected to competition with free unlabeled antigen (Figure 4). Filter lifts from positive plaques were exposed to ^{125}I-labeled BSA-NPN in the presence of increasing concentrations of the inhibitor NPN. A number of phages correlated with NPN-binding as in Figure 3 were spotted in duplicate (about 100 particles per spot) directly onto a bacterial lawn. The plate was then overlaid with an IPTG-soaked filter and incubated for 19 hours at 25°C. The filters were then blocked in 1 percent BSA in PBS before incubation in ^{125}I-labeled-BSA-NPN as done previously with the inclusion of varying amounts of NPN in the labeling solution. Other conditions and procedures were as described for Figure 3. The results for a phage of moderate affinity are shown in duplicate in the figure. Similar results were obtained for four other phages with some differences in the effective inhibitor concentration ranges. These studies showed that individual clones could be distinguished on the basis of antigen affinity. The concentration of free haptens required for complete inhibition of binding varied between 10 to 100 x 10^{-9}M, suggesting that the expressed Fab fragments had binding constants in the nanomolar range.

In preparation for characterization of the protein products, a plasmid containing the heavy and light chain genes was excised with helper phage in an analogous fashion as that for lambda Zap II (Figure 5). Briefly, M13mp8 was used as helper phage and the excised plasmid was infected

into a F^+ derivative of MC1061. The excised plasmid contains the same constructs for antibody fragment expression as do the parent vectors (Figure 1). These plasmid constructs are more conveniently analyzed for restriction pattern and protein expression of the lambda phage clones identified and isolated on the basis of antigen binding. The plasmid also contains an f1 origin of replication which facilitates the preparation of single-stranded DNA for sequence analysis and in vitro mutagenesis. Mapping of the excised plasmid demonstrated a restriction pattern consistent with incorporation of heavy and light chain sequences. The protein products of one of the clones was analyzed by enzyme-linked immunosorbent assay (ELISA) and immunoblotting to establish the composition of the NPN binding protein. A bacterial supernatant after IPTG induction was concentrated and subjected to gel filtration. Fractions in the molecular size range 40 to 60 kD were pooled, concentrated, and subjected to a further gel filtration separation. ELISA analysis of the eluted fractions (Figure 6) indicated that NPN binding was associated with a protein of a molecular size of about 50 kD, which contained both heavy and light chains.

For ELISA characterization, the concentration partially purified bacterial supernatant of an NPN binding clone was separated by gel filtration and samples from each fraction were applied to microtiter plates coated with BSA-NPN. Addition of either antibody to decapeptide (---) or antibody to K chain (-, left-hand scale) conjugated with alkaline phosphatase was followed by color development. The arrow indicates the position of elution of known Fab fragment. The results show that antigen binding is a property of a 50-kD protein containing both heavy and light chains. To permit protein characterization, a single plaque of a NPN-positive clone (Figure 3) was picked, and the plasmid containing the heavy and light chain inserts

(Figure 5) was excised as described above. Cultures (500 ml) in L broth were inoculated with 3 ml of a saturated culture of the clone and incubated for 4 hours at 37°C. Protein synthesis was induced by the addition of IPTG to a final concentration of 1 mM, and the cultures were incubated for 10 hours at 25°C. The supernatant from 200 ml of cells was concentrated to 2 ml and applied to a TSK-G4000 column. Microtiter plates were coated with BSA-NPN at 1 μg/ml, 50 μl samples from the eluted fractions, were mixed with 50 μl of PBS-Tween 20 (0.05 percent) BSA (0.1 percent) added, and the plates were incubated for 2 hours at 25°C. The plated material was then washed with PBS-Tween 20-BSA and 50 μl of appropriate concentrations of a rabbit antibody to decapeptide or a goat antibody to mouse K light chain (Southern Biotech, Oakridge, TN) conjugated with alkaline phosphatase were added and incubated for 2 hours at 25°C. The plates were again washed, 50 μl of p-nitrophenyl phosphate (1 mg/ml in 0.1 tris, pH 9.5, containing 50 mM $MgCl_2$) was added, and the plates were incubated for 15 to 30 minutes and the absorbance was read at 405 nm.

An immunoblot of a concentrated bacterial supernatant preparation under nonreducing conditions was developed with antibody to decapeptide. This revealed a 50-kD protein band. We have found that the antigen-binding protein can be purified to homogeneity from bacterial supernate in two steps involving affinity chromatography on protein G followed by gel filtration. SDS-PAGE analysis of the protein revealed a single band at approximately 50 kD under nonreducing conditions and a doublet at approximately 25 kD under reducing conditions. Taken together, these results are consistent with NPN-binding being a function of Fab fragments in which heavy and light chains are covalently linked by a disulfide bond.

EXAMPLE V

PROPERTIES OF THE IN VIVO REPERTOIRE COMPARED TO THE PHAGE COMBINATORIAL LIBRARY

A moderately restricted library was prepared only because a limited number of primers was used for polymerase chain reaction (PCR) amplification of Fd sequences. The library is expected to contain only clones expressing K-gammal sequences. However, this is not an inherent limitation of the method since the addition of more primers can amplify any antibody class or subclass. Despite this restriction, a large number of clones producing antigen binding proteins were able to be isolated.

A central issue is how the phage library compares with the in vivo antibody repertoire in terms of size, characteristics of diversity, and ease of access.

The size of the mammalian antibody repertoire is difficult to judge, but a figure of the order of 10^6 to 10^8 different antigen specificities is often quoted. With some of the reservations discussed below, a phage library of this size or larger can readily be constructed by a modification of the method described. Once an initial combinatorial library has been constructed, heavy and light chains can be shuffled to obtain libraries of exceptionally large numbers.

In principle, the diversity characteristics of the naive (unimmunized) in vivo repertoire and corresponding phage library are expected to be similar in that both involve a random combination of heavy and light chains. However, different factors act to restrict the diversity expressed by an in vivo repertoire and phage library. For example, a physiological modification such as tolerance will restrict the expression of certain antigenic specificities from the in vivo repertoire, but these

specificities may still appear in the phage library. However, bias in the cloning process may introduce restrictions into the diversity of the phage library. For example, the representation of mRNA for sequences expressed by stimulated B cells can be expected to predominate over those of unstimulated cells because of higher levels of expression. In addition, the resting repertoire might overrepresent spontaneously activated B cells whose immunoglobulins have been suggested to be less specific. I any event, methods exist to selectively exclude such populations of cells. Also, the fortuitous presence of restriction sites in the variable gene similar to those used for cloning and combination will cause them to be eliminated. We can circumvent some of these difficulties by making minor changes, such as introducing amber mutations in the vector system. Different source tissues (for example, peripheral blood, bone marrow, or regional lymph nodes) and different PCR primers (for example, those to amplify different antibody classes), may result in libraries with different diversity characteristics.

Another difference between in vivo repertoire and phage library is that antibodies isolated from the repertoire may have benefited from affinity maturation as a result of somatic mutations after combination of heavy and light chains whereas the phage library randomly combines the matured heavy and light chains. Given a large enough phage library derived from a particular in vivo repertoire, the original matured heavy and light chains will be recombined. However, since one of the potential benefits of this technology is to obviate the need for immunization by the generation of a single highly diverse "generic" phage library, it would be useful to have methods to optimize sequences to compensate for the absence of somatic mutation and clonal selection. Three procedures are made readily available through the vector system presented. First, saturation mutagenesis may be performed

on the complem ntarity-determining regions (CDR's) (23) and the resulting Fab's can be assayed for increased function. Second, a heavy or a light chain of a clone that binds antigen can be recombined with the entire light or heavy chain libraries, respectively, in a procedure identical to that used to construct the combinatorial library. Third, iterative cycles of the two above procedures can be performed to further optimize the affinity or catalytic properties of the immunoglobulin. The last two procedures are not permitted in B cell clonal selection, which suggests that the methods described here may actually increase the ability to identify optimal sequences.

Access is the third area where it is of interest to compare the in vivo antibody repertoire and phage library. In practical terms the phage library is much easier to access. The screening methods used have allowed one to survey the gene products of at least 50,000 clones per plate so that 10^6 to 10^7 antibodies can be readily examined in a day but the most powerful screening methods depend on selection. In the catalytic antibody system, this may be accomplished by incorporating into the antigen leaving groups necessary for replication of auxotrophic bacterial strains or toxic substituents susceptible to catalytic inactivation. Further advantages are related to the fact that the in vivo antibody repertoire can only be accessed via immunization, which is a selection on the basis of binding affinity. The phage library is not similarly restricted. For example, the only general method to identify antibodies with catalytic properties has been by preselection on the basis of affinity of the antibody to a transition state analog. Such restrictions do not apply to the in vitro library where catalysis can, in principle, be assayed directly. The ability to assay directly large numbers of antibodies for function may allow selection for catalysts in reactions where a mechanism is not well defined or synthesis of the transition state analog is

difficult. Assaying for catalysts directly eliminates the bias of the screening procedure for reaction mechanisms limited to a particular synthetic analog; therefore, simultaneous exploration of multiple reaction pathways for a given chemical transformation are possible.

We have described procedures for the generation of Fab fragments that are clearly different in a number of important respects from antibodies. There is undoubtedly a loss of affinity in having monovalent Fab antigen binders, but it is possible to compensate for this by selection of suitably tight binders. For a number of applications such as diagnostics and biosensors, monovalent Fab fragments may be preferable. For applications requiring Fc effector functions, the technology already exists for extending the heavy chain gene and expressing the glycosylated whole antibody in mammalian cells.

The data show that it is now possible to construct and screen at least three orders of magnitude more clones with monospecificity than previously possible. The data also invite speculation concerning the production of antibodies without the use of live animals.

EXAMPLE VI
Flp RECOMBINANCE

The lambda Lc 2 vector was constructed for the cloning of PCR amplified products of mRNA that code for light chain protein, as described in Example II, by inserting the nucleotide sequence shown in Table 3 into the Sac I and Xho I sites of lambda Zap II. The vector was prepared by digesting 10 μg of lambda arms from the Uni-Zap^{TM} XR Vector Kit (Stratagene, La Jolla, CA) with 30 units in 100 μl reaction Sac I. Overlapping synthetic oligonucleotides were cloned into the above Sac I digested arms as follows. Oligonucleotides L11 through L15 and L17 - L19 (L11, L12,

L13, L14, L15, L17, L18 and L19) (shown in Table 3) were kinased by adding 1 μl of each oligonucleotide (0.1 μg/μl) and 20 units of T_4 polynucleotide kinase (BRL, Gaithersburg, MD) to a solution containing 70 mM Tris HCL at pH 7.6, 0.1 M KCl, 10 mM $MgCl_2$, 5 mM DTT, 1 mM adenosine triphosphate (ATP), 10 mM 2 ME, 500 micrograms per ml of BSA. The solution was maintained at 37°C for 30 minutes and the reaction stopped by maintaining the solution at 65°C for 10 minutes. The two end oligonucleotides L16 and L110 were added to the above kinasing reaction solution together with 1/10 volume of a solution containing 20 mM Tris-HCL at pH 7.4, 2.0 mM $MgCl_2$ and 50.0 mM NaCl. This solution was heated to 70°C for 5 minutes and allowed to cool slowly to room temperature. During this time period all oligonucleotides annealed to form the double stranded synthetic DNA insert similar to the one shown in Figure 2A. The annealed oligonucleotides were covalently linked to each other by adding 40 μl of the above reaction to a solution containing 66 mM Tris-HCL at pH 7.6, 6.6 mM $MgCl_2$, 1 mM DTT, 1 mM ATP and 10 units of T4 DNA ligase (BRL, Gaithersburg, MD). This solution was maintained at 25°C for 30 minutes and then the T4 DNA ligase was inactivated by heating the solution at 65°C for 10 minutes. The unphosphorylated ends of the resultant oligonuleotides were kinased by mixing 52 μl of the above reaction, 4 μl of a solution containing 10 mM ATP and 5 units of T4 polynucleotide kinase. This solution was maintained at 37°C for 30 minutes and then the T4 polynucleotide kinase was inactivated by heating the solution at 65°C for 10 minutes. The phosphorylated synthetic DNA insert was ligated directly into the above prepared lambda Zap II vector arms.

TABLE 5

L11	TGAATTCTAAACTAGTCGCCAAGGAGACAG
L12	TCATAATGAAATACCTATTGCCTACGGCAG
L13	CCGCTGGATTGTTATTACTCGCTGCCCAAC
L14	CAGCCATGGCCGAGCTCGTCAGTACTAGTG
L15	TTAAGCGGCCGCAA
L16	TCGATTGCGGCCGCTTAACACTAGTACTGACGA
L17	GCTCGGCCATGGCTGGTTGGGCAGCGAGTA
L18	ATAACAATCCAGCGGCTGCCGTAGGCAATA
L19	GGTATTTCATTATGACTGTCTCCTTGGCGA
L110	CTAGTTTAGAATTCAAGCT

TABLE 6

H11	GGCCGCAAATTCTATTTCAAGGAGACAGTC
H12	ATAATGAAATACCTATTGCCTACGGCAGCC
H13	GCTGGATTGTTATTACTCGCTGCCCAACC
H14	AGCCATGGCCCAGGTGAAACTGCTCGAGA
H15	TTCTAGCTAGTTACCCGTACGACGTTCC
H16	GGACTACGGTTCTTAATAGAATTCG
H17	TCGACGAATTCTATTA
H18	AGAACCGTAGTCCGGAACGTCGTACGGG
H19	TAACTAGACTAGTAATCTCGAGCAGTTTC
H110	ACCTGGGCCATGGCTCCTTGGGCAGCGAGT
H111	AATAACAATCCAGCGGCTGCCGTAGGCAA
H112	TAGGTATTTCATTATGACTGTCTCCTT
H113	GAAATAGAATTTGC

The lambda Hc 2 vector was constructed for cloning PCR amplified products coding for heavy chain Fd sequences, as described in Example II, by inserting the nucleotide sequence shown in Figure 2B into the Not I and Xho I sites of lambda Zap II. As with the light chain vector, the heavy chain vector was prepared by digesting lambda arms from the Uni-Zap™ XR Vector Kit (Stratagene, La Jolla, CA)

with 30 units of Not I restriction enzyme in 100 µl reaction buffer. The inserted sequence similar to the one in Figure 2B was constructed from the overlapping synthetic oligonucleotides depicted in Table H11 to H113 as outlined above.

The sequence of the oligonucleotides described above include elements for construction, expression, and secretion of Fab fragments. These oligonucleotides introduce the asymmetric Not I and Eco RI restriction sites; a leader peptide for the bacterial pel B gene, which has previously been successfully used in E. coli to secrete Fab fragments, Better et al., Science, 240:1041 (1988); Skerra and Pluckthun, Science, 240:1038 (1988), both of which are incorporated herein by reference, a ribosome binding site at the optimal distance for expression of the cloned sequence; cloning sites for either the light or heavy chain PCR product; and, in lambda Hc 2, a decapeptide tag at the carboxyl terminus of the expressed heavy chain protein fragment. The sequence of the decapeptide tag was useful because of the availability of monoclonal antibodies to this peptide that were used for immunoaffinity purification of fusion proteins, Field et al. Mol. Cell Biol., 8:2159 (1988), which is incorporated herein by reference. The vectors were characterized by restriction digest analysis and DNA sequencing, Sanger et al., Proc. Natl. Acad. Sci., USA, 74:5463-5467 (1977), which is incorporated herein by reference and using AMV Reverse Transcriptase ^{35}S-ATP Sequencing Kit (Stratagene, La Jolla, CA).

The lambda LcRF and lambda LcLF were constructed from lambda Lc2 by inserting the oligonucleotides F01 and F02 or F03 and F04 into the EcoRI site of the lambda Lc2 vector. The vector was prepared for ligation by cleaving 10 µg of lambda Lc2 DNA with 30 units of EcoRI restriction enzyme (NEB Beverly Ma.) in 100 µl of reaction buffer at 37°C for

one hour. The solution was heated to 65°C for 30 minutes and then chilled to 30°C. $CaCl_2$ was added to a final concentration of 5 mM and 5 units Heat-Killable (HK) phosphatase (Epicenter, Madison, WI) was added. The reaction was allowed to preceded for 60 minutes at 30°C. The EcoRI digested lambda Lc2 DNA was purified by phenol chloroform extraction and ethanol precipitation. Lambda LcRF was constructed by ligating three molar equivalence of phosporylated oligonucleotides F03 and F04 to 1 μg of EcoRI digested lambda Lc2 in a 10 μl reaction volume at 4°C overnight. A portion (1 μl) of the ligation was packaged with in vitro lambda phage packaging extract and plated on a lawn of XL1-Blue bacteria.

Lambda LcLF was constructed by ligating three molar equivalence of phosphorylated oligonucleotides F01 and F02 to 1 μg of EcoRI digested lambda Lc2 in 10 μl reaction volume at 4°C overnight. A portion (1 μl) of the ligation mixture was packaged with in vitro lambda phage packaging extract and plated on a lawn of XL1-Blue bacteria.

For identification of desired recombinant phage, fresh LB medium (1% tryptone, 0.5% yeast extract, 1% NaCl, pH 7.5) with 10 μg/ml tetracycline was inoculated with XL1-Blue (Stratagene, La Jolla, CA) and shaken overnight at 37°C. Bacteria were pelleted by centrifugation and resuspended in 1/2 volume of 10 mM $MgSO_4$ 1000 pfu of recombinant phage were mixed with 100 μl of the XL1-Blue suspension and incubated at 37°C for twenty minutes. This mixture was quickly dispersed into 7 ml melted and cooled 0.7% top agarose in LB medium, and the slurry was plated on warmed LB plates. The agarose was allowed to solidify at room temperature and then the plates were warmed at 37°C for 10 to 12 hours, then chilled at 4°C for 1 more hour. MSI Magna Nylon membranes (Fisher Scientific, California) were placed directly onto the agarose surfaces of the plates for 1 minute, and then lifted off carefully. The

plaque lifts were immersed in denaturing solution (0.5 N NaOH, 1.5 M NaCl) for 5 minutes, transferred to neutralization solution (1.5 M NaCl, 0.5 M Tris at pH 7.4) for 5 minutes, rinsed in 2X SSC (0.3 M NaCl, 0.3 M sodium citrate, pH 7.0), and air dried on paper towels. Duplicate filters were generated from the same plates as above. All filters were interleaved between sheets of filter paper and baked at 80°C under constant vacuum for 1 hour. The filters were removed from the filter paper and immersed in wash solution (5X SSC, 0.5% SDS, 1 mM EDTA pH 8.0) at 42°C for 1 to 2 hours; bacterial debris was removed by gently wiping each filter with a sponge soaked in the wash solution. The filters were then prehybridized in prehybridization solution (6X SSC, 0.2% Ficoll 400, 0.2% polyvinylpyrrolidone, 0.2% borine serum albumin (Pentax fraction V), 0.5% SDS, 50 mM sodium phosphate pH 6.5, and 250 μg/ml denatured and sheared herring sperm DNA). After 2 to 14 hours at 65°C, the prehybridization solution was decanted, fresh solution was added along with the oligo probe (see below), and the filters were shaken at room temperature for 14 to 24 hours. The filters were washed three times for 30 minutes at 37°C with 6X SSC/0.5% SDS and then autoradiographed. Plaques showing positive hybridization on both filter and duplicate were isolated and checked by restriction mapping as well as sequencing. Oligomeric probe was prepared as followed: 10 μg oligo was mixed in a 50 μl reaction volume with 50 mM Tris 7.4, 10 mM $MgCl_2$, 5 mM DTT, 0.1 mM EDTA pH 8.0, 0.1 mM spermidine, 100 μCi [α-^{32}P] ATP (Amersham, Arlington Heights, IL), and 10 μ T4 polynucleotide kinase (BRL, Gaithersburg, MD). Oligonucleotide LLF was used to identify lambda LcLF and LRF was used to identify lambda LcRF.

For construction of the lambda LcF vector, lambda LcLF and lambda LcRF vector were crossed at the Xba I site as follows. DNA was first purified from each vector. The lambda LcRF was cleaved with Mlu I restriction

endonuclease, the resulting 5' ends were dephosphorylated, and the product was digested with Xba I. This process cleaved the left arm of the vector into several pieces, but the right arm remained intact. The DNA of lambda LcLF was cleaved with Hind III, dephosphorylated, and then cleaved with Xba I; this process destroyed the right arm, but the left arm remained intact. The DNA's so prepared were then mixed and ligated. After ligation, only clones that resulted from combination of a right arm of light chain-containing clones and left arm of heavy chain-containing clones reconstituted a viable phage. After ligation and packaging, the desired vector, lamba LcF was identified as above by sequence analysis.

The lambda HcRF and lambda HcLF were constructed from lambda Hc3 by inserting the oligonucleotides F01 and F02 or F03 and F04 into the EcoRI site of the lambda Hc2 vector. The vector was prepared for ligation by cleaving 10 μg of lambda Hc2 DNA with 30 units of EcoRI restriction enzyme (NEB Beverly Ma.) in 100 μl of reaction buffer at 37°C for one hour. The solution was heated to 65°C for 30 minutes and then chilled to 30°C. $CaCl_2$ was added to a final concentration of 5 mM and 5 units Heat-Killable (HK) phosphatase (Epicenter, Madison, WI) was added. The reaction was allowed to preceded for 60 minutes at 30°C. The EcoRI digested lambda Hc3 DNA was purified by phenol chloroform extraction and ethanol precipitation. Lambda HcRF was constructed by ligating three molar equivalence of phosporylated oligonucleotides F03 and F04 to 1 μg of EcoRI digested lambda Hc2 in a 10 μl reaction volume at 4°C overnight. A portion (1 μl) of the ligation was packaged with <u>in vitro</u> lambda phage packaging extract and plated on a lawn of XL1-Blue bacteria.

Lambda HcLF was constructed by ligating three molar equivalence of phosphorylated oligonucleotides F01 and F02 to 1 μg of EcoRI digested lambda Hc2 in 10 μl reaction

volume at 4°C overnight. A portion (1 μl) f the ligation mixture was packaged with in vitro lambda phage packaging extract and plated on a lawn f XL1-Blue bacteria. Oligonucleotide HLF was used to identify lambda HcLF and LRF was used to identify lambda LcRF by hybridization, as above.

For construction of the lambda HcF vector, lambda HcLF and lambda HcRF vectors were crossed at the Xba I site. DNA from the two vectors was first purified. The lambda HcRF vector DNA was cleaved with Mlu I restriction endonuclease, the resulting 5' ends were dephosphorylated, and the product was digested with Eco RI. This process cleaved the left arm of the vector into several pieces, but the right arm remained intact. The lambda HcLF DNA was cleaved with Hind III, dephosphorylated, and then cleaved with Eco RI; this process destroyed the right arm, but the left arm remained intact. The DNA's so prepared were then mixed and ligated. After ligation, only clones that resulted from combination of a right arm of light chain-containing clones and left arm of heavy chain-containing clones reconstituted a viable phage. After ligation and packaging the desired heavy chain vector was confirmed by sequence analysis.

Libraries were constructed in these vectors as in Example III except the lambda LcF was cleaved with SacI and SpeI instead of SacI and XbaI in preparation for cloning. The light chain PCR inserts prepared by cleaving with SacI and XbaI were compatible with these arms.

TABLE 7

F01	5' AATTCGAAGTTCCTATACTTTCTAGAG 3'
F02	5' AATTCTCTAGAAAGTATAGGAACTTCG 3'
F03	5' AATTCTCTAGAGAATAGGAACTTCGGAATAGGAACTTCG 3'
F04	5' AATTCGAAGTTCCTATTCCGAAGTTCCTATTCTCTAGAG 3'
LLF	5' TTTCTAGAGAATTCTAAA
LRF	5' GGAACTTCGAATTCTAAA
HLF	5' TTTCTAGAGAATTCGTCGA
HRF	5' GGAACTTCGAATTCGTCGA

EXAMPLE VII

A light chain vector was constructed that contains two amber mutations in the left arm. The left arm was from EMBL3A the right arm was from lambda LcF were combined to construct lambda LcFA. DNA was prepared from each of the two parent lambda phage vectors. 10 ug of lambda EMBL3A was digested with HindIII, dephosphorylated and digested with KpnI. 10 ug of lambda LcF was digested with MluI then dephosphorylated. This DNA was then digested with KpnI. One μg of digested DNA from each of the two parent vectors were mixed and ligated. After packaging in vitro, the phage were infected into BB4 E. coli. All phage had the two amber mutations from EMBL and the cloning site from lambda LcF one of these phage was named lambda LcFA.

A heavy chain vector, lambda HcFA, with an amber mutation in the right arm was constructed from lambda zap (Stratagene, La Jolla, CA) in a manner identical to the construction of Lambda HcFA of example VI except Lambda ZapI was digested with Not I and EcoRI and dephosporylated with heat-kill phosphatase and phage vectors were grown on E. coli BB4.

When heavy and light chain libraries were constructed

in these vect rs as in Example VI, combination could be performed by co-infecting the phage libraries into BB4 (Stratagene, La Jolla, CA) cells which had been transformed with plasmid pUC19F, as described in Govinal and Jayaram, Gene, 51:31-41 (1987), which is incorporated herein by reference, and combined phage were selected by platting on MC1061 (ATCC) E. coli. which selected against SupF amber mutations.

Although the invention has been described with reference to the presently preferred embodiment, it should be understood that various modifications can be made without departing from the spirit of the invention. Accordingly, the invention is limited only by the following claims.

I CLAIM:

1. A composition of matter comprising a plurality of procaryotic cells containing diverse combinations of first and second DNA sequences encoding first and second polypeptides that can be expressed and which form heteromeric receptors and at least one of the plurality of procaryotic cells expressing a heteromer exhibiting binding activity towards a preselected molecule.

2. The composition of matter of claim 1 wherein said procaryotic cells are E. coli.

3. The composition of matter of claim 1 wherein the first and second DNA sequences encode functional portions of heteromeric receptors selected from the group consisting of antibodies, T cell receptors, integrins, hormone receptors and transmitter receptors.

4. The composition of matter of claim 3 wherein said first and second DNA sequences encode functional portions of the variable heavy and variable light chains of an antibody.

5. A composition of matter comprising a plurality of procaryotic cells containing various combinations of diverse first and second DNA sequences encoding first and second polypeptides which can associate to form heteromeric receptors exhibiting binding activity towards preselected molecules, said diversity of first DNA sequence being greater than about 100 different sequences and said diversity of said second DNA sequence being greater than about 1000 different sequences.

6. The composition of matter of claim 5 wherein said procaryotic cells are E. coli.

7. The composition of matter of claim 5 wherein the first and second DNA sequences enc de functional portions of heteromeric receptors selected from the group consisting of antibodies, T cell receptors, integrins, hormone receptors and transmitter receptors.

8. The composition of matter of claim 7 wherein said first and second DNA sequences encode functional portions of the variable heavy and variable light chains of an antibody.

9. A kit for the preparation of vectors useful for the coexpression of two or more DNA sequences, comprising two vectors, a first vector having a first combining site on a defined side of a cloning site which defines orientation and a second vector with a second combining site and a cloning site of orientation asymmetric to that of the first vector, wherein one or both of said vectors contains a promoter for expressing polypeptides which form heteromeric receptors encoded by DNA sequences inserted in said cloning sites.

10. The kit of claim 9 wherein said vectors are in a virus.

11. The kit of claim 9 wherein said vectors are plasmid.

12. The kit of claim 9 wherein said DNA sequences encode functional portions selected from the group consisting of antibodies, T cell receptors, integrins, hormone receptors and transmitter receptors.

13. The kit of claim 12 wherein said DNA sequences encode functional portions of the variable heavy and variable light chains of an antibody.

14. The kit of claim 9 wherein said first and second combining sites are selected from the group consisting of EcoRI-EcoRI, and NotI-NotI.

15. The kit of claim 14 wherein the cloning site is selected from the group consisting of XhoI-SpeI, SacI-XbaI, and SacI-SpeI.

16. A vector, capable of expressing a heteromer exhibiting binding activity towards a preselected molecule when combined with a second vector, having a first combining site on a defined side of a cloning site which defines orientation and which can be combined with a second vector with a second combining site and a cloning site of orientation asymmetric to that of the first vector, wherein one or both of said vectors contains a promoter for expressing polypeptides which form heteromers encoded by DNA sequences inserted in said cloning sites.

17. The vector of claim 16 wherein said DNA sequences encode functional portions of heteromeric receptors selected from the group consisting of antibodies, T cell receptors, integrins, hormone receptors and transmitter receptors.

18. The vector of claim 16 wherein said DNA sequences encode functional portions of the variable heavy and variable light chains of an antibody.

19. A cloning system for the coexpression f two DNA sequences encoding polypeptides which associate to form a heteromer, comprising a set of uniform first vectors having a diverse population of first DNA sequences and a set of uniform second vectors having a diverse population of second DNA sequences, said first and second vectors having complementary combining sites so as to allow the operational combination of said first and second DNA sequences.

20. The cloning system of claim 19 wherein said two DNA sequences encode polypeptides which associate to form heteromeric receptors selected from the group consisting of antibodies, T cell receptors, integrins, hormone receptors and transmitter receptors.

21. The cloning system of claim 20 wherein said two DNA sequences encode functional proteins of the variable heavy and variable light chains of an antibody.

22. The cloning system of claim 19 wherein the combining sites are selected from the group consisting of EcoRI-EcoRI and NotI-NotI.

23. A plurality of expression vectors containing a plurality of possible first and second DNA sequences, wherein each of said expression vectors has operationally linked thereon a first DNA sequence and a second DNA sequence, and wherein substantially each of said vectors contains a different combination of first and second DNA sequence.

24. A method of constructing a diverse population of vectors having first and second DNA sequences encoding first and second polypeptides which associate to form heteromeric receptors, comprising the steps of

(a) operationally linking a diverse population of first DNA sequences encoding said first polypeptides to a first vector having a combining site and a cloning site in a defined orientation;

(b) operationally linking a diverse population of second DNA sequences encoding said second polypeptides to a second vector having a combining site compatible with the combining site on said first vector and a cloning site in an asymmetric orientation to that of the first vector;

(c) combining the vector products of step (a) with the vector products of step (b) under conditions to permit their combination into a combined vector having said first and second DNA sequences operationally linked thereon.

25. The method of claim 24 wherein said first and second DNA sequences encode functional portions of heteromeric receptors selected from the group consisting of antibodies, T cell receptors, integrins, hormone receptors and transmitter receptors.

26. The method of claim 25 wherein said first and second DNA sequences encode functional portions of the variable heavy and variable light chains of an antibody.

27. The method of claim 24 wherein said combining is accomplished by restriction endonuclease cleavage of said vectors of step (a) and (b) and combining said cleaved vectors of step (a) and (b) with DNA ligase.

28. The method of claim 24 wherein said combining is accomplished by Flp recombinase.

29. A method of selecting a procaryotic cell which expresses a heteromer specific for a preselected molecule comprising randomly combining first vectors having a diverse population of DNA sequences encoding polypeptides with second vectors having different diverse populations of DNA sequences which encode polypeptides and which form heteromeric receptors with said polypeptides encoded by said first vector, transfecting a sufficient number of said randomly combined sequences into said procaryotic cells, screening said cells to determine the cell expressing a heteromer specific for said preselected molecule.

30. The method of claim 29 wherein said first and second DNA sequences encode functional portions of heteromeric receptors selected from the group consisting of antibodies, T cell receptors, integrins, hormone receptors and transmitter receptors.

31. The method of claim 30 wherein said first and second DNA sequences encode functional portions of the variable heavy and variable light chains of an antibody.

32. The method of claim 29 wherein said combining is accomplished with restriction endonuclease cleavage of said first and second vectors and ligating the cleaved first and second vectors.

33. The method of claim 29 wherein said combining is accomplished with Flp recombinase.

34. The method of claim 29 wherein the number of randomly combined sequences is sufficiently equivalent to the possible combinations of said populations of said first and second DNAs.

35. A method for identifying functional heteromeric receptors composed of a plurality of polypeptides, comprising coexpressing random combinations of first and second DNA homologs which encode polypeptides which associate to form heteromeric receptors so as to form a diverse population of said first and second DNA homologs, said diversity being at least enough that at least one heteromer formed by the polypeptides resulting from said coexpression has a desired functional property and restricted so that said heteromeric receptors can be screened for a predetermined function.

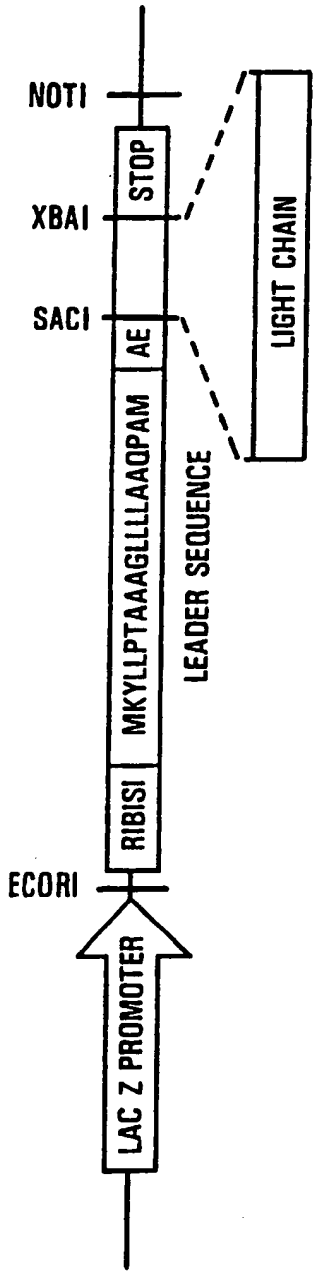

NOTI
XBAI
SACI
ECORI
STOP
AE
RIBISI
MKYLLPTAAAGLLLLAAQPAM
LEADER SEQUENCE
LIGHT CHAIN
LAC Z PROMOTER

FIG. 1-1

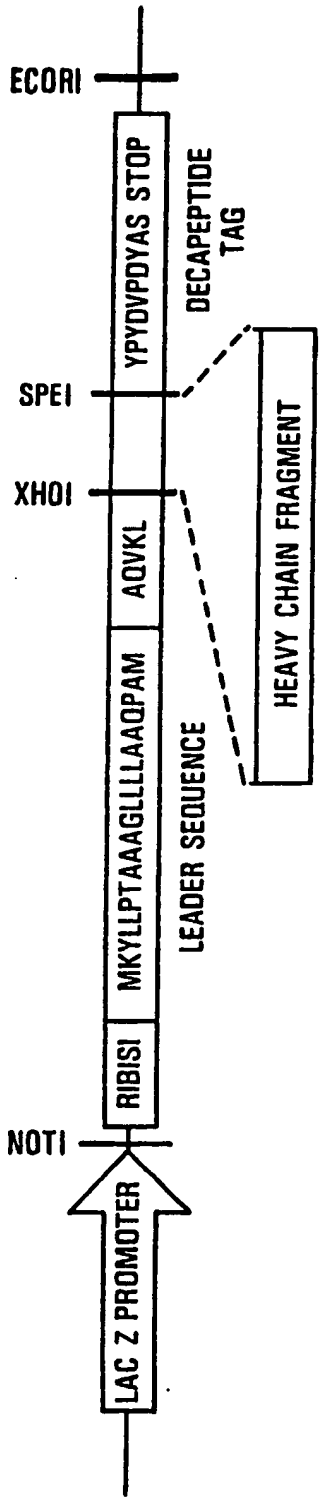

ECORI
SPEI
XHOI
NOTI
YPYDVPDYAS STOP
DECAPEPTIDE TAG
AQVKL
RIBISI
MKYLLPTAAAGLLLLAAQPAM
LEADER SEQUENCE
HEAVY CHAIN FRAGMENT
LAC Z PROMOTER

FIG. 1-2

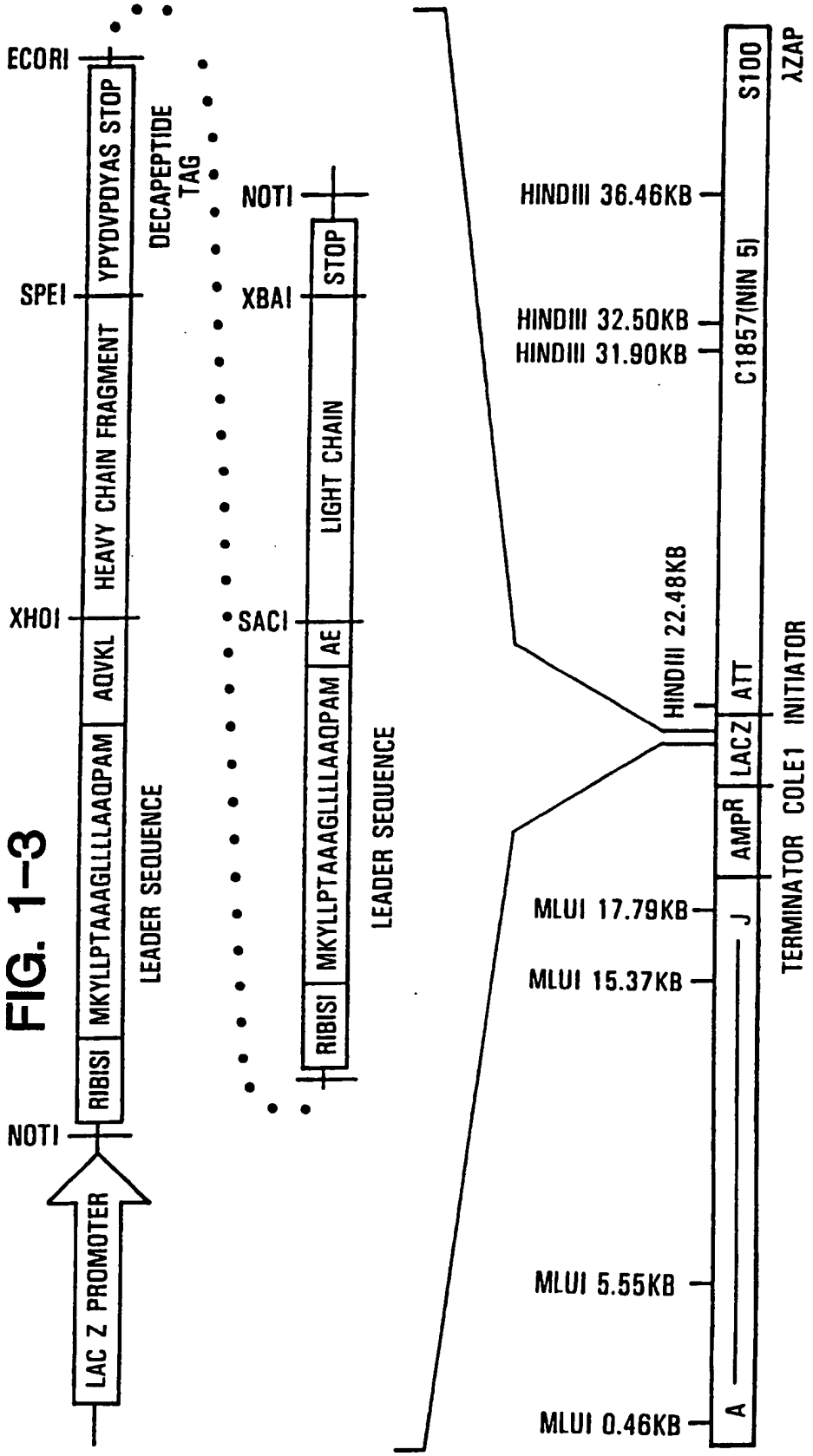

ECORI
SPEI
XHOI
NOTI
LAC Z PROMOTER
RIBISI
MKYLLPTAAAGLLLLAAQPAM
AQVKL
HEAVY CHAIN FRAGMENT
YPYDVPDYAS STOP
DECAPEPTIDE TAG
LEADER SEQUENCE
NOTI
XBAI
SACI
RIBISI
MKYLLPTAAAGLLLLAAQPAM
AE
LIGHT CHAIN
STOP
LEADER SEQUENCE
HINDIII 36.46KB
HINDIII 32.50KB
HINDIII 31.90KB
HINDIII 22.48KB
MLUI 17.79KB
MLUI 15.37KB
MLUI 5.55KB
MLUI 0.46KB
A
J
AMPR
LACZ
ATT
C1857(NIN 5)
S100
λZAP
TERMINATOR COLE1 INITIATOR

FIG. 1–3

FIG. 2A λLc1

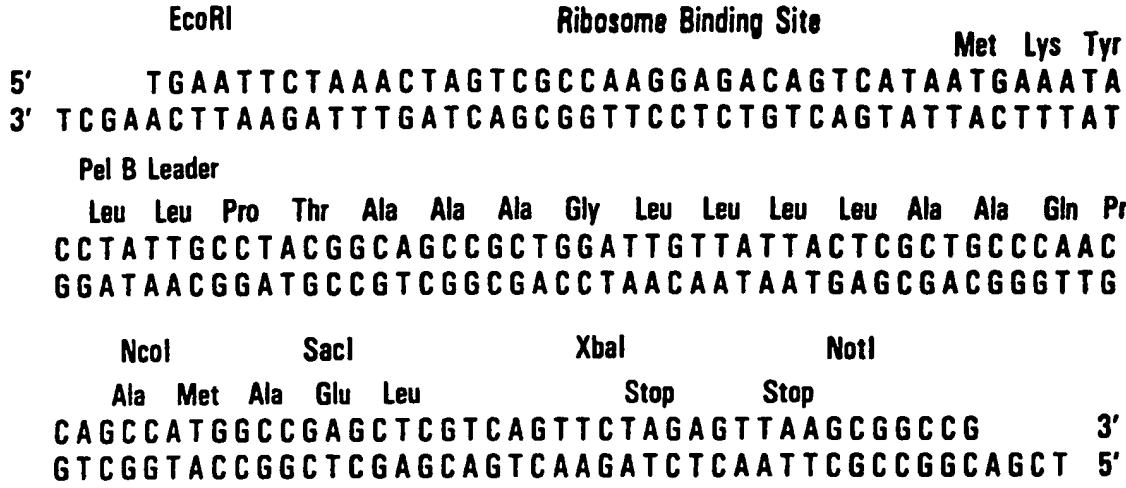


FIG. 2B λHc2

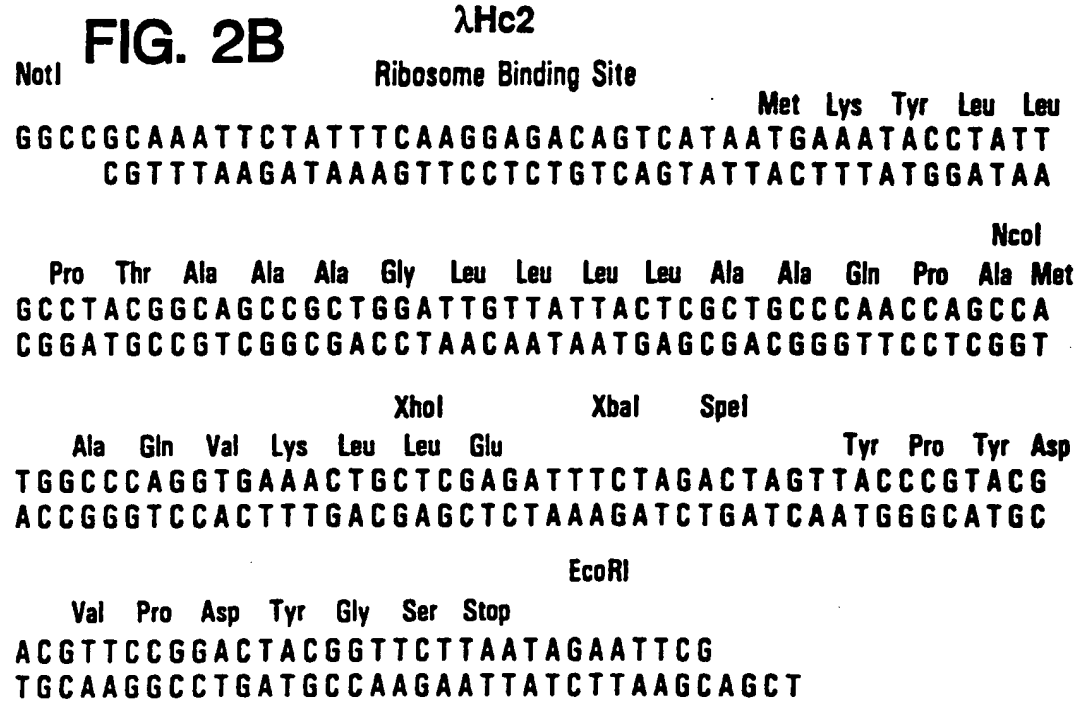


FIG. 3A

FIG. 3B

FIG. 3C

FIG. 3D

FIG. 3E

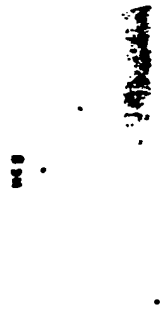


FIG. 3F

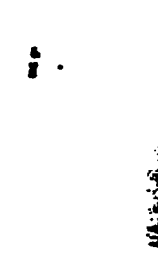


FIG. 3G

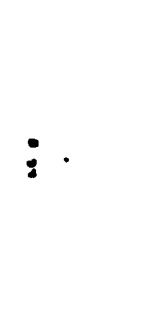


FIG. 3H

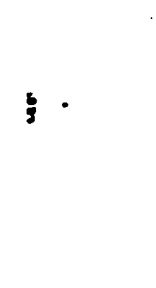


FIG. 4

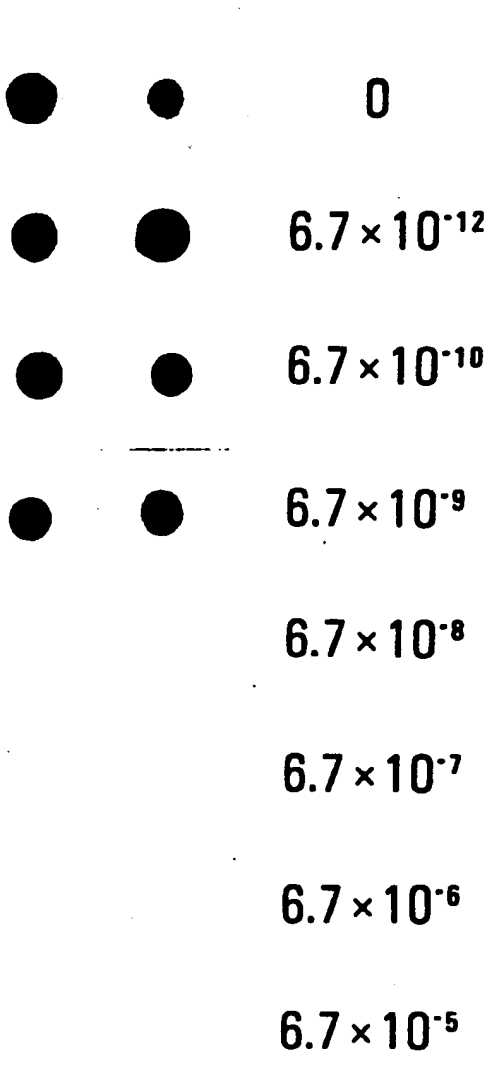


FIG. 5

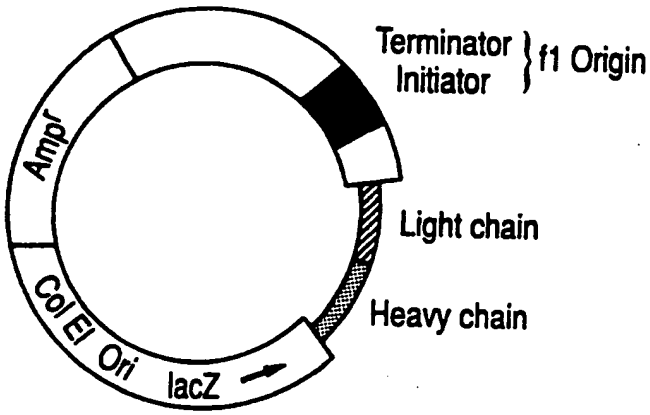

Terminator
Initiator
f1 Origin
Ampr
Light chain
Heavy chain
ColE1 Ori
lacZ

FIG.6

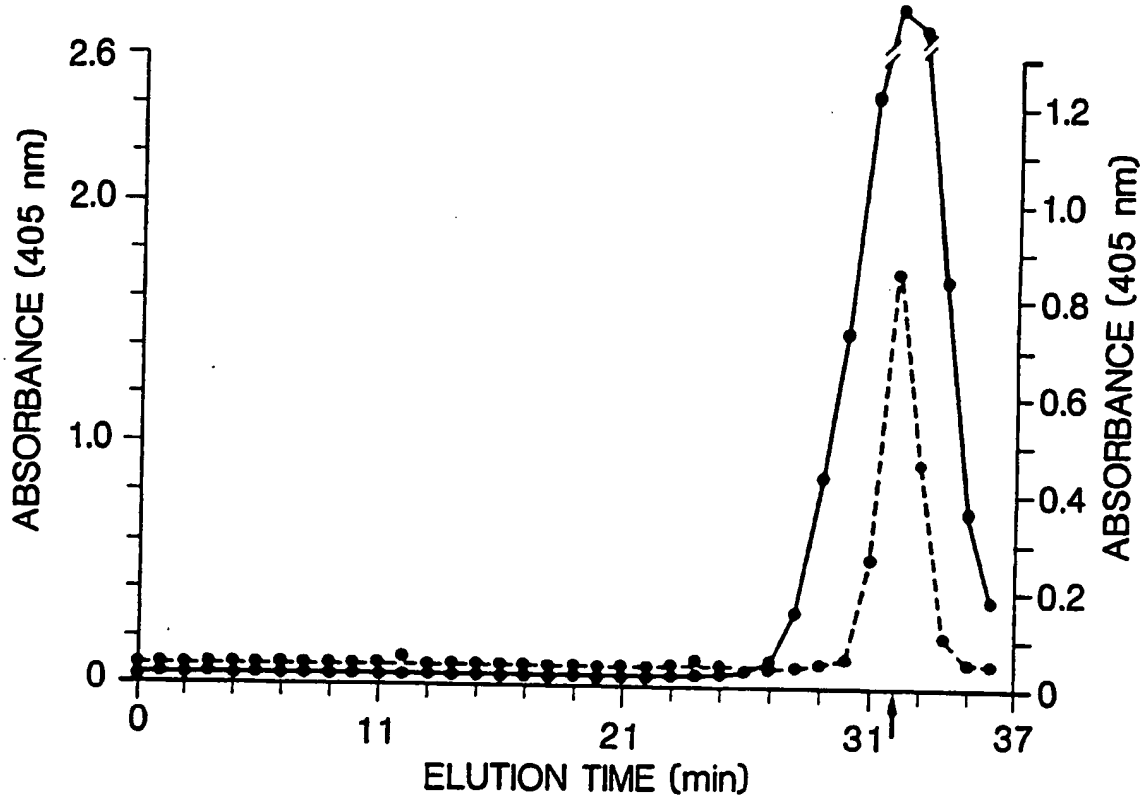
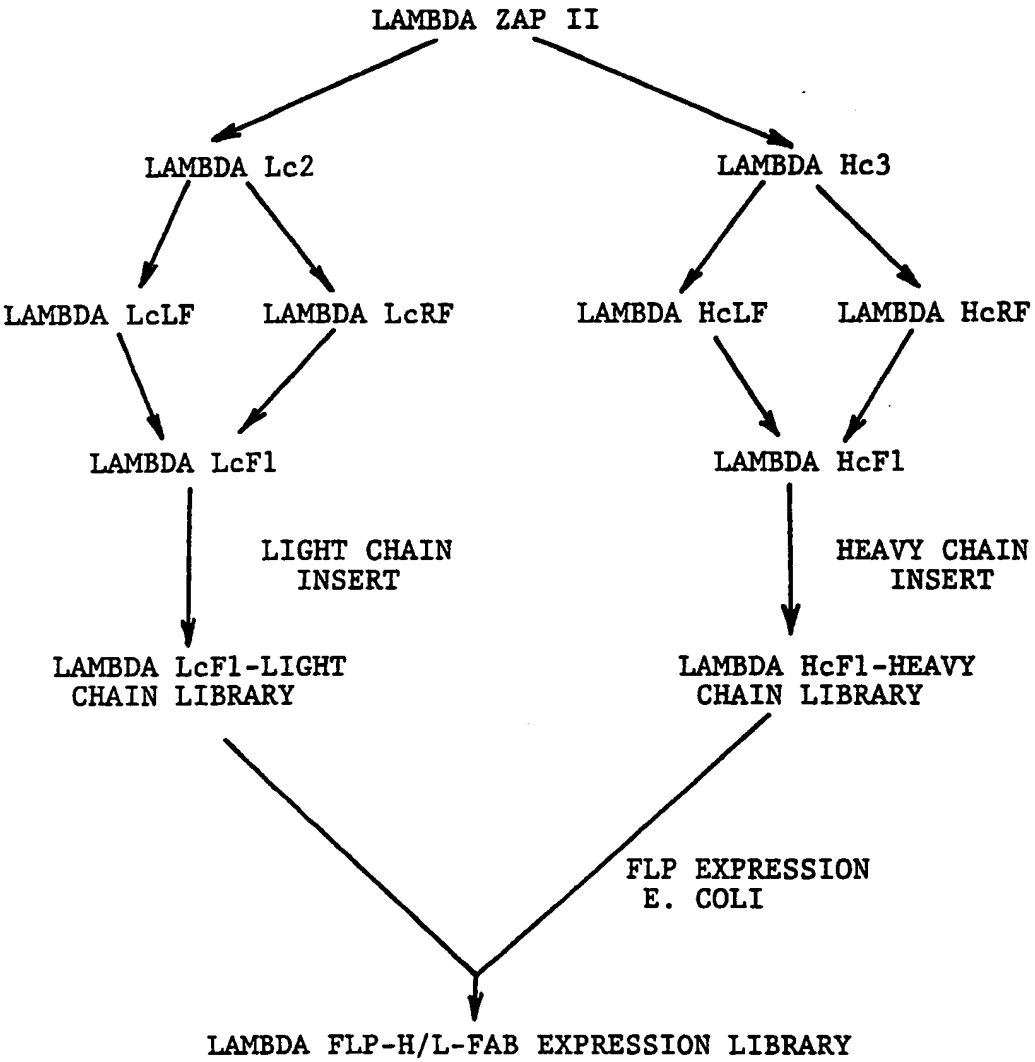

ABSORBANCE (405 nm)
2.6
2.0
1.0
0
1.2
1.0
0.8
0.6
0.4
0.2
0
0
11
21
31
37
ELUTION TIME (min)

FIG. 7


LAMBDA ZAP II
LAMBDA Lc2
LAMBDA Hc3
LAMBDA LcLF
LAMBDA LcRF
LAMBDA HcLF
LAMBDA HcRF
LAMBDA LcF1
LAMBDA HcF1
LIGHT CHAIN INSERT
HEAVY CHAIN INSERT
LAMBDA LcF1-LIGHT CHAIN LIBRARY
LAMBDA HcF1-HEAVY CHAIN LIBRARY
FLP EXPRESSION E. COLI
LAMBDA FLP-H/L-FAB EXPRESSION LIBRARY

INTERNATIONAL SEARCH REPORT

International Application No PCT/US90/02890

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [3]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC(5): C12Q 1/70; C12Q 1/68; C12P 19/34; C12N 15/00; C12Q 1/02;
U.S.CL.: 435/5,6,9,172.3,29,317.1,71.2,320; 935/69.7,80,88

II. FIELDS SEARCHED

Minimum Documentation Searched [4]

Classification System	Classification Symbols
US	435/5,6,91,172.3,29,317.1,71.2,320 935/69.7,80,88

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [5]

APS

III. DOCUMENTS CONSIDERED TO BE RELEVANT [14]

Category *	Citation of Document, [16] with indication, where appropriate, of the relevant passages [17]	Relevant to Claim No. [18]
Y	Nucleic Acids Research, Vol. 16, No. 15, issued 1988, Short et al "λZap expression vector with in vivo excision properties", pages 7583-7600. See Fig. 1A.	10,11,14-16,19, 22,24,27-29, 32-35
Y	Science, Vol. 240, issued 20 May 1988, Skerra et al "Assembly of a Functional Immunoglobulin Fv Fragment in Escherichia coli", pages 1038-1040. See Abstract, Fig. 1, page 1039, col. 3.	1-4,6-8,20,23, 24
X	Stratagene Catalog, issued 1988, pages 16 & 17. See entire document	9-11,14-16,19, 22
Y	US, A, 4,800,159. (Mullis et al) 24 January 1989. See entire document.	5,12,13,17,18, 21,25,26,30,31
Y	US, A, 4,816,397. (Boss et al) 28 March 1989. See entire document.	1-4,6-8,19-22

* Special categories of cited documents: [15]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search [2]	Date of Mailing of this International Search Report [2]
30 August 1990	1 OCT 1990
International Searching Authority [1]	**Signature of Authorized Officer** [20]
ISA/US	Suzanne Ziska

Form PCT/ISA/210 (second sheet) (May 1986)

FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

V. ☐ OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE [1]

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. ☐ Claim numbers . because they relate to subject matter [1] not required to be searched by this Authority, namely:

2. ☐ Claim numbers . . . because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out [1], specifically:

3. ☐ Claim numbers_____, because they are dependent claims not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI. ☒ OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING [2]

This International Searching Authority found multiple inventions in this international application as follows:

I. Claims 1-8 are a composition of matter (cells) comprising a first product; claims 29-35 are a process of using the cells comprising a screening assay.

(SEE ATTACHMENT)

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.

Telephone Practice

2. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4. ☐ As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest

☐ The additional search fees were accompanied by applicant's protest.

☐ No protest accompanied the payment of additional search fees.

Attachment to PCT/ISA/210
PCT/US90/02890

VI. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING "Continuation".

II. Claims 9-15 comprising a kit; a second product.

III. Claims 6-18 & 23, comprising a vector, a third product.

IV. CLaims 19-22, comprising a cloning system (a fourth product)

V. Claims 24-28, comprising a (second) process of use of the cloning system which is the construction of a population of vectors containing different DNA sequences.